

Table des matières

Table des figures	3
1 Généralités et quelques notions probabiliste	12
1.1 Variables aléatoires continues	12
1.2 Fonction W de Lambert	14
1.3 Fonction Quantile	16
1.4 Convergence	17
1.5 Estimation	18
1.5.1 Estimateur sans Biais	19
1.5.2 Erreur quadratique moyenne	19
1.5.3 Construction d'estimateurs	20
1.6 Courbe de Lorenz	21
2 Distribution de Lindley et ses applications	23
2.1 Introduction	23
2.2 Distribution de Lindley	24
2.3 Moments et mesures connexes	25
2.4 Fonction de hasard et fonction de survie	28
2.5 Ordre stochastique	29
2.6 Courbe de Lorenz	30
2.7 Statistiques d'ordre extrêmes	31
2.8 Fonction Quantile de la distribution Lindley	32
2.9 Estimation	33
2.9.1 Estimation par la méthode des moments (MoM)	33
2.9.2 Estimation par la méthode du maximum de vraisemblance	35
2.10 Simulation	36
2.11 Distribution de Gamma Lindley (<i>Gal</i>)	38
2.11.1 Introduction	38
2.11.2 Moments et mesures connexes	39
2.11.3 Courbe de Lorenz	41

2.11.4	Statistiques d'ordres extrêmes	41
2.11.5	Fonction Quantile de la distribution de Gamma Lindley	43
2.11.6	Estimation	43
2.12	Distribution de Size Biased Lindley Distribution et ses propriétés :	45
2.12.1	Fonction de survie et fonction de hasard	47
2.12.2	Moments et mesures connexes	48
3	Distribution de Size Baised Gamma Lindley (SBGaL) immédiate propriétés	51
3.1	Introduction	51
3.2	Distribution Size Baised Gamma-Lindley (SBGaLD) et certaines propriétés	52
3.2.1	Fonction de survie et la La fonction de hasard	53
3.3	Statistiques Moments, courbe de Lorenz et ordre extrême	54
3.3.1	Moments et mesures connexes	54
3.3.2	Courbe de Lorenz	57
3.3.3	Statistiques de commande extrêmes	58
3.4	Estimation des paramètres	59
3.4.1	Estimation du maximum de vraisemblance (ML)	59
3.4.2	Moments estimés	60
3.5	Simulation	61
3.5.1	Application aux ensembles de données réelles	64
	Bibliographie	69

Table des figures

1.1	Fig. 1.1 : <i>Les deux branches de la fonction de Lambert</i>	15
2.1	Fig. 2.1 : <i>Présentation graphique de la fonction de densité pour quelques valeurs de θ, noir ($\theta = 0.5$); rouge ($\theta = 1$); bleu ($\theta = 2$).</i>	25
2.2	Fig. 2.2 : <i>Présentation graphique du coefficient de variation γ (noir), le coefficient de dissymétrie $\sqrt{\beta_1}$ (rouge) et le coefficient d'aplatissement β_2 (bleu)</i>	26
2.3	Fig. 2.3 : <i>Présentation graphique de la fonction de survie pour quelques valeurs de θ, noir ($\theta=0.5$); rouge ($\theta=1$); bleu ($\theta=2$).</i>	29
2.4	Fig. 2.4 : <i>Présentation graphique de la fonction de taux de hasard pour quelques valeurs de θ, noir ($\theta=0.5$); rouge ($\theta=1$); bleu ($\theta=2$).</i>	30
3.1	Fig 3.1. <i>Présentation graphique de la fonction de densité pour quelques valeurs de θ</i>	52
3.2	.2. <i>Présentation graphique de la fonction de répartition pour quelques valeurs de θ</i>	53

Abstract

This thesis proposes a new distribution, called as **A Size Baised Gamma Lindley** distribution (SBGaL). Various statistical properties like stochastic ordering, moment method, maximum likelihood estimation, mean deviations and limiting distribution of extreme order statistics is established. An application of the model to a real data set is presented finally and compared with the fit attained by some other well-known one and two parameters distributions.

Résumé

Dans cette thèse nous introduisons une nouvelle distribution nommée **A Size Biased Gamma Lindley**(SBGaL). Diverses propriétés statistiques telles que l'ordre stochastique, la méthode du moment, l'estimation du maximum de vraisemblance, les écarts moyens et la distribution limite des statistiques d'ordre extrême sont établies. Une application du modèle à un ensemble de données réelles est finalement présentée et comparée à l'ajustement obtenu par d'autres distributions bien connues à un et deux paramètres.



bihbjhjbbwkjvevc

xf

b

fb

f

Introduction

La qualité des procédures utilisées dans une analyse statistique dépend fortement sur le modèle de probabilité supposé ou la distribution. À cause de cela, des efforts considérables ont été déployés dans le développement de grandes classes de distributions de probabilité standard ainsi que des méthodologies statistiques pertinentes. Cependant, il reste beaucoup de problèmes importants où les données réelles ne suivent aucune des modèles de probabilité standard.

Soit X une variable aléatoire suivant la distribution à un paramètre avec la fonction de densité

$$f(x; \theta) = \begin{cases} \frac{\theta^2(1+x)e^{-\theta x}}{1+\theta} & x, \theta > 0 \\ 0, & \text{sinon} \end{cases} \quad (\text{L})$$

Introduit par Lindley (1958). Cette distribution a suscité l'intérêt de nombreux chercheurs et a été généralisée plusieurs fois par divers auteurs. En premier lieu, Sankaran (1970) utilisé (L) lorsque le paramètre suit une loi de Poisson pour dériver leur distribution de Poisson Lindley (PLD) discrète avec fonction de densité

$$f_{PLD}(x; \theta) = \frac{\theta^2(x + \theta + 2)e^{-\theta x}}{(1 + \theta)^x}, \quad x = 0, 1, \dots, \theta > 0.$$

Par ailleurs, certains chercheurs ont proposé de nouvelles classes de distributions basées sur des modifications de la distribution de Lindley, y compris leurs propriétés. L'idée principale est toujours dirigée en intégrant les anciennes distributions à des structures plus flexibles voir (Ghitany et al. (2008, a), Mahmoudi and Zakerzadeh (2010), Dolati (2010), Asgharzadeh et al. (2013), Zeghdoudi et al. (2017,2018,...)).

En raison de l'existence d'un seul paramètre, la distribution de Lindley ne fournit pas suffisamment de flexibilité pour analyser différents types de données à vie. Pour augmenter la flexibilité à des fins de modélisation, il sera utile d'envisager d'autres alternatives de cette distribution.

Les distributions Size Baised, constituent un cas particulier de la forme plus générale connue sous le nom de distributions pondérées. Introduites pour la première fois par Fisher (1934) pour modéliser la détermination biais, les distributions pondérées ont ensuite été formalisées dans une théorie unificatrice par Rao (1965). De telles distributions apparaissent naturellement dans la pratique lorsque les observations d'un échantillon sont enregistrées avec une probabilité inégale, telle que la probabilité proportionnelle à dessins de taille (PPS). En bref, si la variable aléatoire X a la distribution $f(x; \theta)$, avec paramètres inconnus θ , la distribution pondérée correspondante est de la forme

$$f^\alpha(x, \theta) = \frac{w(x)f(x; \theta)}{E(w(X))}$$

où $w(x)$ est une fonction de pondération non négative telle que $E(w(X))$ existe.

Un cas particulier des distributions pondérées, les distributions biaisées par la taille est proposé par Rao (1965) lorsque la fonction pondérée est de la forme $w(x) = x^\beta$, appelée distribution d'ordre biaisé par la taille β , où $\beta = 1$ ou $\beta = 2$, appelé longueur. respectivement à base de zones et à zones biaisées. Par conséquent, la densité de la distribution biaisée en longueur est définie par

$$f_L = \frac{xf(x, \theta)}{E(X)}, \quad -\infty < x < \infty. \quad (0.1)$$

Les distributions pondérées ont de nombreuses applications, nous tenons à citer quelques-uns des résultats les plus récents sur les size biased.

Warren (1975) a été le premier à les appliquer dans le cadre de l'échantillonnage de cellules de bois. Van Deusen (1986) est arrivé indépendamment à la théorie de la distribution biaisée par la taille et l'appliquer à la distribution appropriée des données de diamètre à hauteur d'homme (DHP) issues d'inventaires par échantillonnage horizontal (HPS) (Grosenbaugh, 1958).

Par la suite, Lappi et Bailey (1987) ont utilisé des distributions pondérées pour analyser les données d'incrément de diamètre HPS. Plus récemment, des distributions pondérées ont été utilisées par Magnussen et Al. (1999) pour récupérer la distribution des hauteurs de la canopée à partir de mesures au scanner laser. En écologie, Dennis et Patil (1984) utilisent des méthodes stochastiques. équations différentielles pour arriver à une distribution gamma pondérée en tant que fonction de densité de probabilité stationnaire (PDF) pour un modèle de population stochastique avec des effets de prédation. Dans les pêcheries, Taillie et Al. (1995) ont modélisé des populations de stocks de poissons utilisant des distributions pondérées. Récemment, Arooj Ayesha (2017) a introduit et étudié une nouvelle distribution nommée Size Baised Lindly avec densité de probabilité

$$f_{SBL}(x, \theta) = \frac{\theta^2(1+x)e^{-\theta x}(\theta+x+2)}{1+\theta}, \quad x > 0.$$

Zeghdoudi et Nedjar (2016a, 2016b) ont récemment introduit une nouvelle distribution, nommée distribution Gamma Lindley, basée sur des mélanges de distributions gamma $(2, \theta)$ et de Lindley à un paramètre. où la fonction de densité de la variable aléatoire X est donnée par :

$$f_{GaL}(x; \theta, \beta) = \frac{\theta^2((\beta + \beta\theta - \theta)x + 1)e^{-\theta x}}{\beta(1 + \theta)} \quad x > 0, \theta > 0, \beta > \frac{\theta}{1+\theta}.$$

Dans cette thèse, nous introduisons une nouvelle distribution de durée de vie en utilisant les modèles pondérés (Size Baised) donné par la relation (0.1).

Notre projet de recherche est structuré de la manière suivante :

Le premier chapitre est consacré aux rappels des certaines définitions et certains résultats qui nous utilisons par la suite à savoir : généralités et quelques distributions de probabilités, fonction W de Lambert, fonction quantile, statistiques d'ordre extrêmes, estimation MM et MV, courbe de Lorenz.

Dans le chapitre II, nous faisons une synthèse sur la distribution de **size biased Lindley** en nous inspirant des travaux de Lindley (1958), Ghitany et al. (2008, a) et Arooj Ayesha (2017).

Le chapitre III se focalise sur la nouvelle distribution à deux paramètres (size biased gamma Lindley) dont on donne quelques propriétés à savoir : courbe de Lorenz, méthode des moments, estimation du maximum de vraisemblance et la distribution de limitation des statistiques d'ordre extrême. Plusieurs simulations sont établies pour examiner le biais et l'erreur quadratique moyenne des estimateurs des paramètres ob-

tenus par la méthode du maximum de vraisemblance et leur application dans l'analyse de survie. Par ailleurs, une comparaison des distributions est obtenue.

Chapitre 1

Généralités et quelques notions probabiliste

Dans cette introduction on présente les caractéristiques essentielles des données à analyser ainsi que les outils techniques permettant de décrire leur distribution.

1.1 Variables aléatoires continues

Si on considère la variable aléatoire $X(X \geq 0)$, représentant la durée de vie, alors sa loi de probabilité peut être définie par l'une des cinq fonctions équivalentes suivantes (chacune des fonctions ci-dessous peut être obtenue à partir de l'une des fonctions).

Fonction de répartition

Soit F la fonction de répartition de la variable X , s'il existe une fonction f

positive, intégrable telle que :

$$F(x) = \mathbf{P}(T \leq x) = \int_0^x f(t) dt \quad (1.1)$$

Fonction de densité

C'est la fonction $f \geq 0$ telle que,

$$\int_0^\infty f(x) dx = 1 \quad (1.2)$$

Fonction de survie

La fonction de survie pour x fixé est définie comme :

$$S(x) = \mathbf{P}(X > x) = 1 - F(x), x \geq 0 \quad (1.3)$$

Sachant que S décroissante, continue à droite avec $S(0) = 1$ et $\lim_{t \rightarrow \infty} S(t) = 0$.

Risque instantané h (ou taux de hasard)

La fonction de risque instantané est la fonction

$$h(t) = \frac{f(t)}{S(t)} \quad (1.4)$$

Si cette fonction n'est pas une densité de probabilité on peut définir le taux de hasard comme suit :

$$h(t) = \lim_{\Delta x \rightarrow 0} \frac{\mathbf{P}[x \leq X < x + \Delta x | X \geq x]}{\Delta x} \quad (1.5)$$

Taux de hasard cumulé H

La fonction de risque cumulé est donnée par :

$$H(x) = \int_0^x h(s)ds \quad (1.6)$$

Avec l'égalité suivante entre fonction de survie et fonction de risque cumulé :

$$H(x) = -\ln S(x) \quad (1.7)$$

Le temps moyen et la variance de la durée de survie

Le temps moyen de survie $E(X)$ et la variance de la durée de survie $V(X)$ sont définis par les quantités suivantes :

$$\mathbf{E}(X) = \int_0^\infty S(x)dt \quad (1.8)$$

$$\mathbf{Var}(T) = 2 \int_0^\infty xS(x)dt - \left(\int_0^\infty S(x)dt \right)^2 \quad (1.9)$$

Ainsi on peut déduire l'espérance et la variance à partir de n'importe laquelle des fonctions F, S, f, H, h (mais pas l'inverse).

1.2 Fonction W de Lambert

La fonction Lambert W est définie comme l'inverse à plusieurs valeurs de la fonction $w \rightarrow we^w$. Elle a de nombreuses applications en mathématiques pures et appliquées.

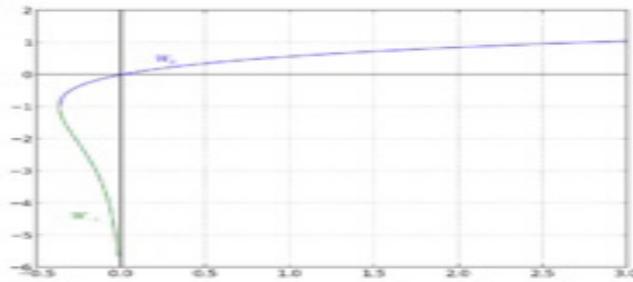


FIG. 1.1: **Fig. 1.1** : *Les deux branches de la fonction de Lambert*

La fonction W de Lambert (Lambert 1758) est une fonction complexe multivaluée définie comme étant la solution de l'équation :

$$W(z) \exp(W(z)) = z. \quad (1.10)$$

Où z est un nombre complexe. Si z est un nombre réel tel que $z \geq -1/e$ alors $W(z)$ devient une fonction réelle et il y a deux possibles branches réelles.

La branche réelle prenant des valeurs dans $]-\infty, -1]$ est appelée la branche négative et notée W_{-1} .

La branche réelle prenant des valeurs dans $[-1, \infty[$ est appelée la branche principale et notée W_0 . Tous les deux réelles branches de W sont représentés sur la figure.1.1.

En plus de l'équation. (1.5), L'émeray [14] a indiqué que d'autres équations peuvent être résolues en termes de la fonction W de Lambert. À cet égard, le lemme suivant, sera essentiel dans notre thèse.matière

Lemme 1.1 (*L'émeray et al.* [14])

Soient a, b et c des nombres complexes fixés. La solution de l'équation $z + ab^z = c$ par rapport à $z \in \mathbb{C}$ est :

$$z = c - \frac{1}{\log(b)} W(ab^c \log(b)).$$

Où W représente la fonction W de Lambert.

Preuve. Pour tous nombres complexes a, b et c fixés, on doit résoudre l'équation $z + ab^z = c$ par rapport à la variable complexe z . Multiplier par $b^c \log(b)$ les deux côtés de cette équation, l'équation résultante peut être écrite comme suit :

$$(c - z) \log(b) \exp((c - z) \log(b)) = ab^c \log(b). \quad (1.11)$$

On considère maintenant l'équation (1.16) conjointement avec l'équation (1.15). Il est clair que $(c - z) \log(b)$ est la fonction W de Lambert de l'argument complexe $ab^c \log(b)$. Par conséquent, on a

$$W(ab^c \log(b)) = (c - z) \log(b).$$

Ce qui implique le résultat souhaité. Ceci termine la preuve du lemme 1.1.

1.3 Fonction Quantile

Soit X une variable aléatoire et F sa fonction de répartition, la fonction quantile est définie par

$$Q_X(u) = F^{\leftarrow}(u) = \inf \{x \text{ tel que } F_X(x) \geq u.\} \quad 0 < u < 1. \quad (1.12)$$

Pour toute valeur de $u \in]0, 1[$, la notion F^{\leftarrow} désignant l'inverse généralisé à gauche de F .

1.4 Convergence

Définition 1.1

Une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires est dite convergente en probabilité vers la variable aléatoire X si :

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) \rightarrow 0, \forall \varepsilon > 0. \quad (1.13)$$

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables de fonction de répartition F_n et X une variable aléatoire de fonction de répartition F . On dit que la suite de variable aléatoire $(X_n)_{n \in \mathbb{N}}$ converge en loi vers la variable aléatoire X ou en distribution si

$$\lim_{n \rightarrow \infty} F_n(x) = F(x) \quad (1.14)$$

On dit que $(X_n)_{n \in \mathbb{N}}$ converge presque sûrement vers X si

$$P(\lim_{n \rightarrow \infty} X_n = X) = 1 \quad (1.15)$$

La suite $(X_n)_{n \in \mathbb{N}}$ converge en moyenne quadratique vers la variable aléatoire X si

$$\lim_{n \rightarrow \infty} E[(X_n - X)^2] = 0 \quad (1.16)$$

Théorème 1.1.(*Théorème central limite*)

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires indépendantes et de même loi de probabilité, d'espérance mathématique m et de variance σ^2). on pose

$$Z_n = \frac{\bar{X}_n - m}{\frac{\sigma}{\sqrt{n}}} \rightsquigarrow N(0, 1). \quad (1.17)$$

Alors, quand $n \rightarrow +\infty$, Z_n tend en loi vers une variable normale centrée réduite.

La variable aléatoire Z_n converge en loi vers la loi normale centrée réduite. C'est une conséquence du TCL qui assure que :

$$\sqrt{n} (\bar{X}_n - m) \xrightarrow{p} N(0, \sigma^2).$$

1.5 Estimation

Soit (X_1, X_2, \dots, X_n) un échantillon prélevé d'une population de la probabilité F_θ dépendant de un ou plusieurs paramètres θ . Il s'agit d'estimer θ sur la base de l'observation d'un échantillon prélevé de cette population.

Définition 1.2

Un estimateur d'un paramètre inconnu θ est une statistique dont la valeur est une estimation du paramètre θ . l'estimateur d'un paramètre d'une population donnée par un évalueur unique est dite estimation ponctuelle du paramètre.

1.5.1 Estimateur sans Biais

Définition 1.3

Un estimateur $\hat{\theta} = f(x_1, x_2, \dots, x_n)$ de θ est dit sans biais si $E(\hat{\theta}) = \theta$ (où θ est le paramètre inconnu à estimer)

$$\text{Biais}(\hat{\theta}) = E(\hat{\theta}) - \theta. \quad (1.18)$$

est appelé le biais de l'estimateur où $\hat{\theta}$ est estimateur quelconque de θ , Lorsque $\text{Biais}(\hat{\theta}) = 0$, l'estimateur est dit sans biais, et si $\text{Biais}(\hat{\theta}) > 0$, l'estimateur est dit positivement biaisé.

1.5.2 Erreur quadratique moyenne

L'erreur quadratique moyenne (*Mean Squared Error en anglais*) appelée aussi risque quadratique est l'espérance du carré de l'erreur entre la vraie valeur et sa valeur estimée.

$$MSE(\hat{\theta}_n) = E\left(\left(\hat{\theta}_n - \theta\right)^2\right). \quad (1.19)$$

Le risque quadratique MSE permet de mesurer l'erreur faite en estimant θ par $\hat{\theta}_n$.

$\hat{\theta}_n$ est un bon estimateur de θ si $\text{Biais}(\hat{\theta}_n)$ est nul et si $MSE(\hat{\theta}_n)$ est petit. Entre deux estimateurs sans biais de θ , on choisira l'estimateur possédant le plus petit risque quadratique.

Si le risque est faible, l'estimateur $\hat{\theta}$ est proche de θ . L'erreur quadratique moyenne

1.5.3 Construction d'estimateurs

Si l'on cherche à évaluer un paramètre inconnu θ d'une loi de probabilité P_θ , on dispose de différentes méthodes pour construire un estimateur : Méthode des Moments, Méthode du maximum de vraisemblance...etc.

Méthode des Moments

Elle consiste à estimer les paramètres recherchés en égalisant certains moments théoriques(espérance mathématique, variance etc...) (qui dépendent de ces paramètres) avec leurs contreparties empiriques(moyenne empirique, variance empirique etc...)

Autrement dit, si $\theta = E(X)$, alors l'estimateur de θ par la méthode des moments est

$$\hat{\theta}_n = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (1.20)$$

Plus généralement, pour $\theta \in \Theta$, si $E(X) = \varphi(\theta)$, où φ est une fonction inversible, alors l'estimateur de θ par la méthode des moments est :

$$\hat{\theta}_n = \varphi^{-1}(\bar{X}).$$

De la même manière, on estime la variance de la loi des X_i par la variance empirique de l'échantillon $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$.

Méthode du maximum de vraisemblance

Définition 1.4

Soit (X_1, X_2, \dots, X_n) un n -échantillon de loi P_θ , Soit f_θ la densité de probabilité associée à P_θ . La fonction de vraisemblance de l'échantillon (X_1, X_2, \dots, X_n) s'écrit :

$$L(\theta; x_1, \dots, x_n) = \prod_{i=1}^n f_X(x_i; \theta) \quad (1.21)$$

Définition 1.5

On appelle estimateur du maximum de vraisemblance, la variable aléatoire réelle $\hat{\theta}_n$ qui maximise la fonction de vraisemblance. L'estimateur de maximum de vraisemblance (MV) de θ est la variable aléatoire correspondante. Donc $\hat{\theta}_n$ sera en général calculé en maximisant la log-vraisemblance :

$$\hat{\theta}_n = \operatorname{argmax} \log L(\theta; x_1, \dots, x_n).$$

Quand $\theta = (\theta_1, \dots, \theta_d) \in \Theta$ et que toutes les dérivées partielles ci-dessous existent, $\hat{\theta}_n$ est solution du système d'équations appelées équations de vraisemblance :

$$\forall j \in \{1, \dots, d\}, \quad \frac{\partial}{\partial \theta_j} \log L(\theta; x_1, \dots, x_n) = 0,$$

$$\text{Où} \quad \frac{\partial^2}{\partial \theta_j^2} \log L(\theta; x_1, \dots, x_n) < 0.$$

1.6 Courbe de Lorenz

La courbe de Lorenz est une représentation graphique qui met en relation la fraction " $x\%$ " d'une population détentrice d'une part d'une grandeur (richesse), à la

part "y%" de la grandeur détenue. Elle a été développée par Max O. Lorenz en vue d'une représentation des inégalités de revenu. Dans le cas de l'analyse des revenus, la courbe de Lorenz $L(P)$ représente la part du revenu total détenue par la proportion $P \in [0, 1]$ d'individus les plus pauvres :

$$L(P) = \frac{\text{revenu total des plus pauvres}}{\text{revenu total}}.$$

La courbe de Lorenz pour une variable aléatoire X positive est définie comme le graphe du rapport

$$L(P) = \frac{\int_0^x tf(t)}{\int_0^{+\infty} tf(t)} \quad 0 \leq P \leq 1$$

$$\text{Où } P = F(x)$$

$$L(P) = \frac{E(X|X \leq x)P(X \leq x)}{E(X)}.$$

$$L(P) = \frac{E(X|X \leq x)P}{E(X)}. \quad (1.24)$$

Avec les propriétés $L(p) \leq p$, $L(0) = 0$ et $L(1) = 1$. Si X représente le revenu annuel, $L(p)$ est la proportion du revenu total qui revient aux personnes ayant les revenus les plus faibles de 100p%. Si tous les individus gagnent le même revenu alors $L(p) = p$ pour tout p . La zone située entre la ligne $L(p) = p$ et la courbe de Lorenz peut être considérée comme une mesure de l'inégalité des revenus, ou plus généralement, de la variabilité de X , voir Gail et Gastwirth [20] et Dagum [2] pour une vaste discussion des courbes de Lorenz.

Chapitre 2

Distribution de Lindley et ses applications

2.1 Introduction

Dans cette section , on propose la distribution de **Lindley** à un seul paramètre et étudier ses propriétés de base. Cette distribution est introduit par **Lindley** en 1958 comme combinaison d'Exp(θ) et de Gamma ($2, \theta$). Plus de détails sur la distribution de **Lindley** peuvent être trouvés dans les travaux Ghitany et d'autres chercheurs ([16], [17]) .

2.2 Distribution de Lindley

Soit $Y_1 \sim \exp(\theta)$ et $Y_2 \sim \text{Gamma}(2, \theta)$ deux variables aléatoires indépendantes.

On considère la variable aléatoire $X = Y_1$ et $X = Y_2$ avec les probabilités respectivement $P_1 = \frac{\theta}{1+\theta}$ et $P_2 = \frac{1}{1+\theta}$ sachant que $\theta > 0$.

La fonction de densité de X est donnée par :

$$f(x; \theta) = \frac{\theta^2(1+x)e^{-\theta x}}{1+\theta}; \quad x, \theta > 0 \quad (2.1)$$

La fonction de répartition correspondante est :

$$F(x) = 1 - \frac{\theta + 1 + \theta x}{1 + \theta} e^{-\theta x}; \quad x > 0, \theta > 0 \quad (2.2)$$

La première dérivée de (2.1) est :

$$\frac{d}{dx} f(x) = \frac{\theta^2}{1+\theta} (1 - \theta - \theta x) e^{-\theta x}.$$

Il en résulte que

(i) pour $\theta < 1$, $\frac{d}{dx} f(x) = 0$ implique que $x_0 = \frac{1-\theta}{\theta}$ est le point critique unique à laquelle $f(x)$ est maximisée.

(ii) pour $\theta \geq 1$, $\frac{d}{dx} f(x) \leq 0$, c-à-d $f(x)$ diminue en x .

Par conséquent, le mode de cette distribution est :

$$\text{Mode}(X) = \begin{cases} \frac{1-\theta}{\theta}, & 0 < \theta < 1 \\ 0, & \text{ailleurs} \end{cases}$$

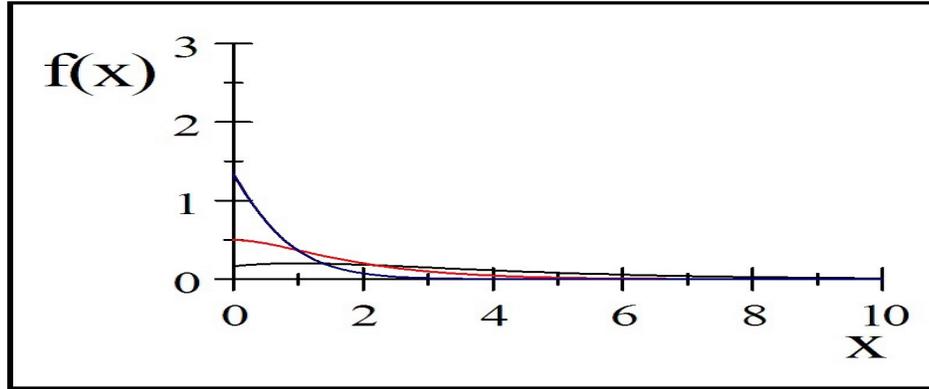


FIG. 2.1: **Fig. 2.1** : Présentation graphique de la fonction de densité pour quelques valeurs de θ , noir ($\theta = 0.5$) ; rouge ($\theta = 1$) ; bleu ($\theta = 2$).

2.3 Moments et mesures connexes

Le moment d'ordre k de la distribution de **Lindley** est :

$$\mu_k^l = E(X^k) = \frac{k! (\theta + k + 1)}{\theta^k (\theta + 1)}, k = 1, 2, \dots$$

d'où, on a :

$$\mu_1^l = \frac{(\theta + 2)}{\theta(\theta + 1)}, \mu_2^l = \frac{2(\theta + 3)}{\theta^2(\theta + 1)}, \mu_3^l = \frac{6(\theta + 4)}{\theta^3(\theta + 1)}, \mu_4^l = \frac{24(\theta + 5)}{\theta^4(\theta + 1)}.$$

Le moment centré d'ordre k de la distribution de **Lindley** est défini par :

$$\mu_k = \mathbf{E}\{(X - \mu)^r\} = \sum_{r=0}^k \binom{k}{r} \mu_k^l (-\mu)^{k-r}.$$

D'où, on a

$$\mu_2 = \frac{\theta^2 + 4\theta + 2}{\theta^2(\theta + 1)} = \sigma^2, \mu_3 = \frac{2(\theta^3 + 6\theta^2 + 6\theta + 2)}{\theta^3(\theta + 1)^3}, \mu_4 = \frac{3(3\theta^4 + 24\theta^3 + 44\theta^2 + 32\theta + 8)}{\theta^4(\theta + 1)^4}.$$

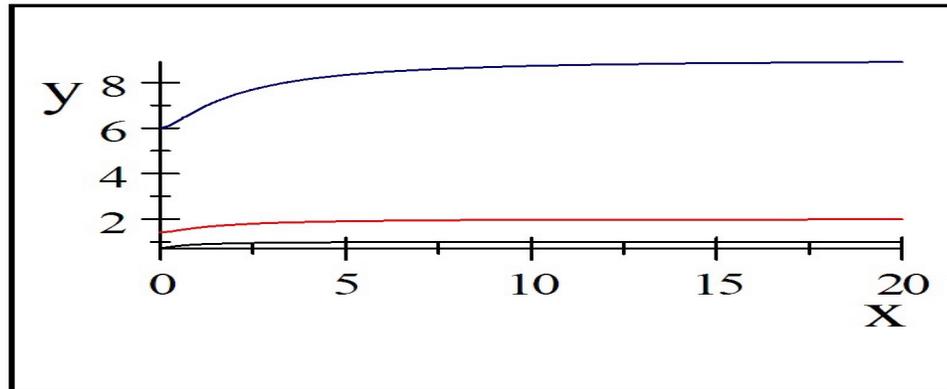


FIG. 2.2: **Fig. 2.2** : Présentation graphique du coefficient de variation γ (noir), le coefficient de dissymétrie $\sqrt{\beta_1}$ (rouge) et le coefficient d'aplatissement β_2 (bleu) .

Le coefficient de variation (γ), le coefficient de dissymétrie *Skewness* ($\sqrt{\beta_1}$) et le coefficient d'aplatissement *Kurtosis* (β_2) sont :

$$\begin{aligned}\gamma &= \frac{\sqrt{\theta^2 + 4\theta + 2}}{\theta + 2}, \\ \text{Skewness}(\sqrt{\beta_1}) &= \frac{2(\theta^3 + 6\theta^2 + 6\theta + 2)}{(\theta^2 + 4\theta + 2)^{\frac{3}{2}}}, \\ \text{Kurtosis}(\beta_2) &= \frac{3(3\theta^4 + 24\theta^3 + 44\theta^2 + 32\theta + 8)}{(\theta^2 + 4\theta + 2)^2}.\end{aligned}$$

Tableau 1.1. *Skewness, kurtosis et coefficient de variation pour certaines valeurs du paramètre θ .*

θ	Skewness pour LD	Kurtosis pour LD	CV pour LD
0.1	1.42249	6.024845	0.7392464
0.5	1.512281	6.342561	0.8246211
0.9	1.600732	6.710286	0.8730337
1.3	1.670306	7.032192	0.9035183
1.7	1.723992	7.299966	0.9240714
2.1	1.765821	7.520627	0.9386284
2.5	1.798906	7.702946	0.9493337
2.9	1.825478	7.854595	0.9574452
3.3	1.847123	7.981736	0.9637429

Remarque 2.1 :

- (i) γ est une fonction croissante en θ et $(1/\sqrt{2}) < \gamma < 1$; voir Fig2.2
- (ii) $\sqrt{\beta_1}$ est une fonction croissante en θ et $\sqrt{2} < \sqrt{\beta_1} < 2$; voir Fig2 : 2.
- (iii) β_2 est une fonction croissante en θ et $6 < \beta_2 < 9$; voir Fig2 : 2.

Théorème 2.1

Soit $X \sim \text{Lindley}(\theta)$. Alors

$$\text{Mode}(X) < \text{Median}(X) < E(X).$$

Preuve. Soient $M = \text{Mode}(X)$, $m = \text{Median}(X)$ et $\mu = \mathbf{E}(X) = \frac{\theta+2}{\theta(\theta+1)}$. Depuis la

fonction de répartition de la distribution de **Lindley**, il en résulte que :

$$F(M) = \begin{cases} 1 - \frac{2}{1+\theta}e^{-(1-\theta)}, & 0 < \theta < 1 \\ 0 & \text{ailleurs} \end{cases}, \quad F(me) = \frac{1}{2}$$

et

$$F(\mu) = 1 - \frac{\theta^2 + 3\theta + 3}{(1+\theta)^2}e^{-\frac{\theta+2}{\theta+1}}.$$

Notons que $F(M)$ est une fonction décroissante en $\theta \in (0, 1)$ et, pour tout $\theta > 0$,

$$0 \leq F(M) < 1 - 2e^{-1} < (1/2).$$

De même, $F(\mu)$ est une fonction croissante en $\theta > 0$ et $(1/2) < 1 - 3e^{-2} < F(\mu) < 1$.

Enfin, étant donné que $F(x)$ est une fonction croissante en $x > 0$. Il est facile de vérifier que $F(M) < F(me) = \frac{1}{2} < F(\mu)$, alors on a $M < m < \mu$.

2.4 Fonction de hasard et fonction de survie

Soit

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x \mid X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} = \frac{\theta^2(1+x)}{\theta + 1 + \theta x}.$$

et

$$S(x) = 1 - F(x) = \frac{\theta + 1 + \theta x}{1 + \theta}e^{-\theta x}.$$

La fonction de taux de hasard et la fonction de survie de la distribution de **Lindley**, respectivement.

Remarque 2.2

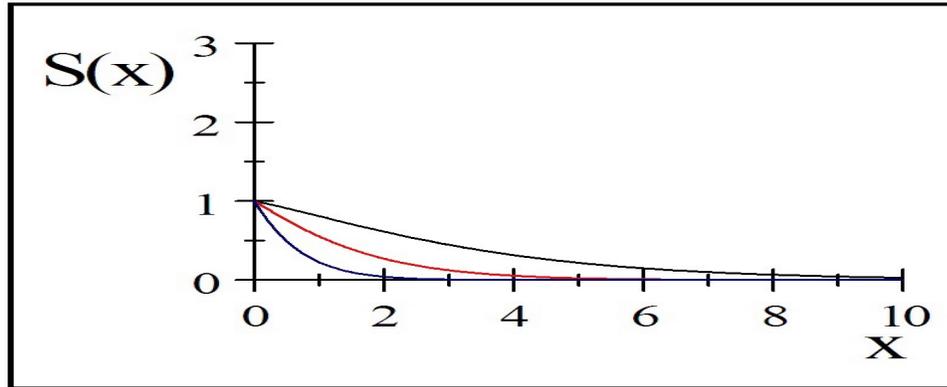


FIG. 2.3: **Fig. 2.3** : Présentation graphique de la fonction de survie pour quelques valeurs de θ , noir ($\theta=0.5$); rouge ($\theta=1$); bleu ($\theta=2$).

$$(i) h(0) = f(0) = \frac{\theta^2}{\theta+1}.$$

(ii) Comme $\frac{d}{dx}h(x) = \frac{\theta^2(1+x)}{\theta+1+\theta x}$, $h(x)$ est une fonction croissante en x et θ en plus

$$\frac{\theta^2}{\theta+1} < h(x) < \theta.$$

2.5 Ordre stochastique

Théorème 1.5

Soient $X \sim \text{Lindley}(\theta_1)$ et $Y \sim \text{Lindley}(\theta_2)$: Si $\theta_1 > \theta_2$ alors $X \preceq_{lr} Y$ et donc

$X \preceq_{hr} Y$ et $X \preceq_s Y$.

Preuve. Notons d'abord que

$$\frac{f_X(t)}{f_Y(t)} = \frac{\theta_1^2(1+\theta_2)}{\theta_2^2(1+\theta_1)} e^{-(\theta_1-\theta_2)t}, \quad t > 0$$

On a, pour $\theta_1 > \theta_2$,

$$\frac{d}{dt} \frac{f_X(t)}{f_Y(t)} = (\theta_2 - \theta_1) \frac{f_X(t)}{f_Y(t)} < 0,$$

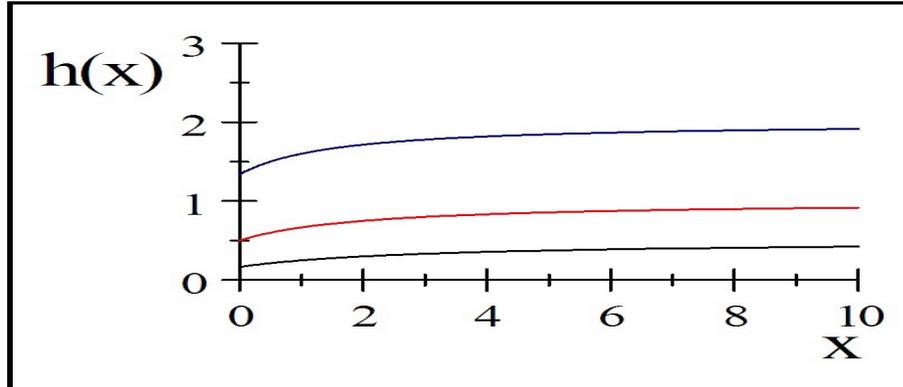


FIG. 2.4: **Fig. 2.4** : Présentation graphique de la fonction de taux de hasard pour quelques valeurs de θ , noir($\theta=0.5$); rouge ($\theta=1$); bleu ($\theta=2$).

$\frac{f_X(t)}{f_Y(t)}$ est décroissante en X . Alors $X \preceq_{lr} Y$. Les états restants découlent des implications dans (1.18).

2.6 Courbe de Lorenz

La courbe de Lorenz pour une variable aléatoire X positif est défini comme le graphe du rapport

$$L(F(x)) = \frac{\mathbf{E}(X|X \leq x)F(x)}{\mathbf{E}(X)} \quad (2.3)$$

Pour la distribution de **Lindley** (1.9) on a,

$$\mathbf{E}(X|X \leq x)F(x) = \frac{2 + \theta}{\theta(1 + \theta)} - \frac{e^{-\theta x}}{1 + \theta} \left[\frac{2}{\theta} + 1 + \theta x^2 + 2x + x\theta \right].$$

Ainsi, à partir de (1.9), on obtient la courbe de Lorenz pour la distribution de **Lindley** comme suit :

$$L(p) = 1 - \frac{\theta(1 + \theta)(1 - p)}{(2 + \theta)(1 + \theta + \theta x)} \left[\frac{2}{\theta} + 1 + \theta x^2 + 2x + x\theta \right].$$

Où $x = F^{-1}(p)$ avec $F(\cdot)$ donnée par (2.2).

2.7 Statistiques d'ordre extrêmes

Si X_1, \dots, X_n un échantillon de n variables aléatoires qui suivent la distribution de **Lindley** et si $\bar{X} = (X_1 + \dots + X_n)/n$ représente la moyenne d'échantillon alors par le théorème central limite $\frac{\sqrt{n}(\bar{X} - \mathbf{E}(X))}{\sqrt{\text{Var}(X)}}$ se rapproche de la distribution normale standard quand $n \rightarrow \infty$.

Théorème 2.3

Parfois, on serait intéressé à étudier la loi asymptotique des valeurs extrêmes $M_n = \max(X_1, \dots, X_n)$ et $m_n = \min(X_1, \dots, X_n)$, Pour la fonction de répartition définie dans (2.2), on constate que :

$$\lim_{t \rightarrow \infty} \frac{1 - F(t+x)}{1 - F(t)} = \exp(-\theta x).$$

et

$$\lim_{t \rightarrow 0} \frac{F(tx)}{F(t)} = x.$$

Ainsi, il en résulte du Théorème 1.4. (Leadbetter et al. [21]) qu'il doit y avoir les constantes de normalisation $a_n > 0, b_n, c_n > 0$ et d_n de telle sorte que :

$$\mathbf{P}\{a_n(M_n - b_n) \leq x\} \rightarrow -\exp\{(-\theta x)\}.$$

et

$$\mathbf{P}\{c_n(m_n - d_n) \leq x\} \rightarrow 1 - \exp(-x). \quad (2.4)$$

Comme $n \rightarrow \infty$. La forme des constantes de normalisation peut également être déterminée. Par exemple, en utilisant le Corollaire 1.1 (Leadbetter et al. [21]), on peut voir que $a_n = 1$ et $b_n = F^{-1}(1 - 1/n)$ avec $F(\cdot)$ donnée par (2.2)

2.8 Fonction Quantile de la distribution Lindley

D'après la fonction de répartition de la distribution **Lindley** définie en (2.2). Il convient de noter qu'est continue et strictement croissante de sorte que la fonction de quantile X est $Q_X(u) = F_X^{-1}(u)$, $0 < u < 1$. Dans le résultat suivant, on donne une expression explicite de Q_X en fonction de la fonction W de **Lambert**.

Théorème 2.4

Pour tout $\theta > 0$, la fonction quantile de la distribution Lindley X est :

$$Q_X(u) = -1 - \frac{1}{\theta} - \frac{1}{\theta} W_{-1} \left(\frac{\theta + 1}{\exp(\theta + 1)} (u - 1) \right), \quad 0 < u < 1, \quad (2.5)$$

Où W_{-1} désigne la branche négative de la fonction W de Lambert.

Preuve. Pour tout θ fixe, $\theta > 0$, soit $u \in (0, 1)$. On doit résoudre l'équation $F_X(x) = u$ par rapport à x , pour tous $X > 0$ comme suit :

$$(\theta + 1 + \theta x) e^{-\theta x} = (\theta + 1) (1 - u). \quad (2.6)$$

En multipliant par $-\exp(-\theta - 1)$ l'équation (2.6), on obtient :

$$-(\theta + 1 + \theta x) \exp(-\theta - 1 - \theta x) = (\theta + 1) (u - 1) \exp(-\theta - 1). \quad (2.7)$$

D'après l'équation (2.7), conjointement avec l'équation (2.5), on voit que $-(\theta + 1 + \theta x)$ est la fonction W de Lambert de l'argument réel $(\theta + 1)(u - 1)\exp(-\theta - 1)$. Alors, on a

$$W\left(\frac{\theta + 1}{\exp(\theta + 1)}(u - 1)\right) = -(\theta + 1 + \theta x), 0 < u < 1. \quad (2.8)$$

Toujours, pour tout $\theta > 0$ et $x > 0$ il est immédiat que $\theta + 1 + \theta x > 1$ et il peut également être vérifié que puisque $u \in (0, 1)$. Il pour, en prenant en compte les propriétés de la branche négative de la fonction W de Lambert a présenté en premier chapitre, l'équation (1.15) devient

$$W_{-1}\left(\frac{\theta + 1}{\exp(\theta + 1)}(u - 1)\right) = -(\theta + 1 + \theta x). \quad (2.9)$$

Ce qui implique le résultat. la preuve du théorème est terminée.

2.9 Estimation

2.9.1 Estimation par la méthode des moments (MoM)

Étant donné un échantillon aléatoire X_1, \dots, X_n , de la distribution de **Lindley** (2.1), l'estimateur des moments (MoM) de θ est :

$$\hat{\theta}_{MoM} = \frac{-(\bar{X} - 1) + \sqrt{(\bar{X} - 1)^2 + 8\bar{X}}}{2\bar{X}}, \bar{X} > 0. \quad (2.10)$$

Le théorème suivant montre que l'estimateur $\hat{\theta}_{MoM}$ de θ est biaisé.

Théorème 2.5

L'estimateur $\hat{\theta}_{MoM}$ de θ est positivement biaisée, i.e. $E(\hat{\theta}) - \theta > 0$.

Preuve. Soient $\hat{\theta}_{MoM} = g(\bar{X})$ et $g(t) = \frac{-(t-1) + \sqrt{(t-1)^2 + 8t}}{2t}$, $\forall t > 0$.

Comme $g''(t) = \frac{1}{t^3} \left[1 + \frac{3t^3 + 15t^2 + 9t + 1}{[(t-1)^2 + 8t]^{\frac{3}{2}}} \right] > 0$, $g(t)$ est strictement convexe.

Ainsi, par l'inégalité de **Jensen**, on a $\mathbf{E}(g(\bar{X})) > g[\mathbf{E}(\bar{X})]$. Enfin, étant donné que

$$\mathbf{E}(g(\bar{X})) = g(\mu) = g\left(\frac{\theta + 2}{\theta(\theta + 1)}\right) = \theta, \quad (2.11)$$

On obtient

$$\mathbf{E}(\hat{\theta}_{MoM}) > \theta.$$

Le théorème suivant donne la loi limite de $\hat{\theta}_{MoM}$.

Théorème 2.6

L'estimateur $\hat{\theta}_{MoM}$ de θ est convergent et asymptotiquement normal :

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{P} N\left(0, \frac{1}{\sigma^2}\right). \quad (2.12)$$

L'intervalle de confiance de θ pour un seuil de confiance $100(1 - \alpha)\%$ est donné par :

$$\hat{\theta} \pm z_{\frac{\alpha}{2}} \frac{1}{\sqrt{n\hat{\sigma}^2}}. \quad (2.13)$$

Où $z_{\frac{\alpha}{2}}$ est le $(1 - \frac{\alpha}{2})$ percentile de la distribution normale standard.

Preuve. Étant donné μ est finie, $\bar{X} \xrightarrow{P} \mu$. $g(t)$ est une fonction continue à $t = \mu$,

$g(\bar{X}) \xrightarrow{P} g(\mu)$, c'est-à-dire $\hat{\theta} \xrightarrow{P} \theta$. Comme $\sigma^2 < \infty$, par le théorème central limite,

on a

$$\sqrt{n}(\bar{X} - \mu) \xrightarrow{P} N(0, \sigma^2). \quad (2.14)$$

En outre, puisque $g(\mu)$ est différentiable et $g'(\mu) \neq 0$, par la méthode Delta, on a :

$$\sqrt{n} (g(\bar{X}) - g(\mu)) \xrightarrow{\mathbf{P}} N\left(0, [g'(\mu)]^2 \sigma^2\right). \quad (2.15)$$

Enfin, étant donné que

$$g(\bar{X}) = \hat{\theta}_{MoM}, g(\mu) = \theta, \quad \text{et} \quad g'(\mu) = \frac{-1}{2\mu^2} \left[1 + \frac{1+3\mu}{\sqrt{(\mu-1)^2 + 8\mu}} \right] = -\frac{1}{\sigma^2}.$$

2.9.2 Estimation par la méthode du maximum de vraisemblance

Soient $X_i \sim LD(\theta)$, $i = \overline{1, n}$ n variables aléatoires. La fonction de logvraisemblance est :

$$\ln l(x_i; \theta) = 2n \ln \theta - n \ln(\theta + 1) + \sum_{i=1}^n \ln(x_i + 1) - n\theta \bar{X}. \quad (2.16)$$

L'estimateurs de la méthode du maximum de vraisemblance $\hat{\theta}_{MV}$ de θ est une solution de l'équation :

$$\frac{\partial \ln l(x_i; \beta, \theta)}{\partial \theta} = \frac{2n}{\theta} - \bar{X} - \frac{n}{(\theta + 1)} = 0.$$

On obtient

$$\hat{\theta}_{MV} = \frac{-(\bar{X} - 1) + \sqrt{(\bar{X} - 1)^2 + 8\bar{X}}}{2\bar{X}}, \bar{X} > 0.$$

Avec

$$\frac{\partial^2 \ln l(x_i; \beta, \theta)}{\partial \theta^2} = -\frac{2n}{\theta^2} - \sum_{i=1}^n \frac{x_i^2}{(\theta x_i + \theta)^2} < 0.$$

Remarque 2.3

L'estimateur de la méthode des moments $\hat{\theta}_{MoM}$ et l'estimateur du maximum de vraisemblance $\hat{\theta}_{MV}$ du paramètre θ sont les mêmes.

2.10 Simulation

Cette section étudie le comportement des estimateurs du maximum de vraisemblance $\hat{\theta}_{MV}$ pour un échantillon de taille finie (n). La simulation est réalisée pour chaque couple $(\theta; n)$, où $\theta = 0.1, 1, 9$ et $n = 20, 40, 60, 80, 100$. Alors on a l'algorithme suivant :

- Choisir les valeurs initiales de θ_0 pour spécifier la distribution de **Lindley** ;
- Choisir la taille de l'échantillon n ;
- Générer N échantillons indépendants de taille n de $LD(\theta)$;
- Calculer les estimations $\hat{\theta}_{MV}$ de θ pour chacun des N échantillons ;
- Calculer :

(i) La moyenne des estimateurs obtenus sur tous les N échantillons

$$\text{biais moyen}(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\Theta}_i - \Theta_0).$$

(ii) L'erreur quadratique moyenne EQM des estimations simulées

$$EQM(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\Theta}_i - \Theta_0)^2. \quad (2.17)$$

Tableau 2.1 : les biais moyens de l'estimateur $\hat{\theta}$

n	$\theta = 0.1$	$\theta = 1$	$\theta = 9$
20	0.00265	0.03068	0.41459
40	0.00118	0.01699	0.18019
60	0.00090	0.01060	0.13665
80	0.00061	0.00746	0.09073
100	0.00058	0.00570	0.07728

Tableau 2.2 : la moyenne EQM de l'estimateur $\hat{\theta}$.

n	$\theta = 0.1$	$\theta = 1$	$\theta = 9$
20	0.00028	0.03335	4.25425
40	0.00013	0.01543	1.91203
60	0.00009	0.01024	0.22017
80	0.00007	0.00752	0.90044
100	0.00005	0.00590	0.70490

Remarque 2.4

(i) Le **tableau 2 :1** présente un biais positif, comme indiqué dans le théorème 2 :3

Le tableau montre également que le biais diminue (augmente) quand $n(\theta)$ augmente

respectivement.

(ii) Le **tableau 2 :2** montre que l'erreur quadratique moyenne diminue (augmente)

lorsque $n(\theta)$ augmente respectivement .

2.11 Distribution de Gamma Lindley (*Gal*)

2.11.1 Introduction

Dans cette subsection, on propose la distribution de Gamma **Lindley**. Cette distribution est introduit par H. **Zeghdoudi**, S. **Nedjar** en (2016)([9], [11]) comme un mélange de Gamma $(2, \theta)$ et de **Lindley** (θ) .

Soit $Y_1 \sim \text{Gamma}(2, \theta)$ et $Y_2 \sim \text{LD}(\theta)$ deux variables aléatoires indépendantes. Pour $\theta > 0$, $\beta > \frac{\theta}{1+\theta}$ les paramètres d'échelle et de mélange respectivement , on considère la variable aléatoire $X = Y_1$ et $X = Y_2$ avec les probabilités respectivement $P_1 = \frac{\beta-1}{\beta}$ et $P_2 = \frac{1}{\beta}$.

La fonction de densité de X est donnée par :

$$f_{Gal}(x; \theta, \beta) = \frac{\theta^2((\beta + \beta\theta - \theta)x + 1)e^{-\theta x}}{\beta(1 + \theta)} \quad x > 0, \theta > 0, \beta > \frac{\theta}{1+\theta}. \quad (2.18)$$

Remarque 2.5

Si $\beta = 1$, cette distribution est la distribution de Lindley $\text{LD}(\theta)$.

Par conséquent, le mode de cette distribution est :

$$\text{Mode}(X) = \begin{cases} \frac{\beta\theta + \beta - 2\theta}{\theta(\beta + \beta\theta - \theta)}, & \forall \beta \in \left[\frac{2\theta}{\theta + 1}, \infty \right] \\ 0, & \text{ailleurs.} \end{cases} \quad (2.19)$$

On peut facilement trouver la fonction de répartition de GaL :

$$F_{\text{GaL}}(x) = 1 - \frac{((\theta\beta + \beta - \theta)(\theta x + 1) + \theta)e^{-\theta x}}{\beta(1 + \theta)}; x > 0, \theta > 0, \beta > \frac{\theta}{1 + \theta}. \quad (2.20)$$

Pour plus détail voir (Zeghdoudi et Nedjar [9]).

2.11.2 Moments et mesures connexes

Proposition 2.1

Soient X_1, X_2, \dots, X_n n variables aléatoires indépendantes de $\text{GaL}(\theta, \beta)$. La fonction génératrice de $S = \sum_{i=1}^n X_i$, est donnée par :

$$M_S(t) = \left(\frac{\theta}{\theta - t} \right)^{2n} \left(\frac{\theta\beta + \beta - t}{\beta(1 + \theta)} \right)^n. \quad (2.21)$$

d'où

$$M_X(t) = \mathbf{E}(e^{tX}) = \left(\frac{\theta}{\theta - t} \right)^2 \left(\frac{\theta\beta + \beta - t}{\beta(\theta + 1)} \right). \quad (2.22)$$

Remarque 2.6

La fonction génératrice des moments pour X et S existent ($\mathbf{E}(e^{tX}) < \infty$) si seulement si $t < \theta$.

Corollaire 2.1

Soit $X \sim GaL(\theta, \beta)$ le moment d'ordre 1, 2 et la variance de X sont :

$$\begin{aligned} \mathbf{E}(X) &= \frac{2\beta(1+\theta) - \theta}{\theta\beta(1+\theta)}, \mathbf{E}(X^2) = \frac{6\beta\theta + 6\beta - 4\theta}{\beta\theta^2 + \beta\theta^3}, \\ \mathbf{Var}(X) &= \frac{-(-2\beta\theta + \theta)^2 + (2 + 6\theta)\beta^2 - 2\beta(\beta\theta - 3\beta\theta^2 + 2\theta^2)}{\beta^2\theta^2(1+\theta)^2}. \end{aligned} \quad (2.23)$$

Théorème 2.6

Soit $X \sim GaLD(\theta, \beta)$, $M = Mode(X)$, $me = Median(X)$ et $Moyenne(X) = \mu = E(X)$ Alors

$$M < me < \mu. \quad (2.24)$$

Preuve. Voir Théorème 1 (**Zeghdoudi et Nedjar (2016)**)[9].

Fonction de hasard et la La fonction de survie

Soit

$$h_{GaL}(x) = \frac{f_{GaL}(x)}{1 - F_{GaL}(x)} = \frac{((\beta + \theta\beta - \theta)x + 1)\theta^2}{\theta(\beta + \theta\beta - \theta)x + \beta + \theta\beta}.$$

et

$$S_{GaL}(x) = 1 - F_{GaL}(x) = \frac{((\theta\beta + \beta - \theta)(\theta x + 1) + \theta)e^{-\theta x}}{\beta(1 + \theta)}$$

La fonction de taux de hasard et la fonction de survie de la distribution de Gamma

Lindley,

respectivement.

Proposition 2.2

Soit $h_{GaL}(x)$ la fonction de taux de risque de X . Alors $h_{GaL}(x)$ est croissante.

Preuve. D'après *Glaser (1980)* et à partir de la fonction de densité (2.18)

$$\rho(x) = -\frac{f'_{GaL}(x)}{f_{GaL}(x)} = -\frac{(\beta - 2\theta + \theta\beta + x\theta^2 - x\theta\beta - x\theta^2\beta)}{x(\beta - \theta + \theta\beta) + 1}.$$

Comme $\rho'(x) = \frac{(\beta - \theta + \theta\beta)^2}{(x\beta - x\theta + x\theta\beta + 1)^2} \geq 0, \forall x, \beta, \theta$ alors $h_{GaL}(x)$ est croissante.

Ordres Stochastiques

Théorème 2.7

Soient $X_i \sim GaL(\theta_i, \beta_i), i = 1, 2$ deux variables aléatoires.

Si $\theta_1 = \theta_2$ et $\beta_1 \geq \beta_2$, alors $X_1 \prec_{lr} X_2, X_1 \prec_{hr} X_2, X_1 \prec_s X_2$ et $X_1 \leq_{cx} X_2$.

Preuve. On a

$$\frac{f_{X_1}(t)}{f_{X_2}(t)} = \frac{\theta_1^2 \beta_2 (1 + \theta_2) ((\beta_1 + \beta_1 \theta_1 - \theta_1) t + 1)}{\theta_2^2 \beta_1 (1 + \theta_1) ((\beta_2 + \beta_2 \theta_2 - \theta_2) t + 1)} e^{-(\theta_1 - \theta_2)t}.$$

Pour simplifier, on utilise $\ln \left(\frac{f_{X_1}(t)}{f_{X_2}(t)} \right)$, on peut trouver

$$\frac{d}{dt} \ln \left(\frac{f_{X_1}(t)}{f_{X_2}(t)} \right) = -(\theta_1 - \theta_2) + \frac{\beta_1 \theta_1 - \beta_2 \theta_2 + (\beta_1 - \beta_2) + (\theta_2 - \theta_1)}{(t\beta_1 \theta_1 + t\beta_1 - t\theta_1 + 1)(t\beta_2 \theta_2 + t\beta_2 - t\theta_2 + 1)}.$$

À cet effet, si $\theta_1 \geq \theta_2$ et $\beta_1 \leq \beta_2$, on a $\frac{d}{dt} \ln \left(\frac{f_{X_1}(t)}{f_{X_2}(t)} \right) \leq 0$. Les états restants découlent des implications dans (1.8).

2.11.3 Courbe de Lorenz

La courbe de Lorenz pour la distribution de Gamma **Lindley**

$$L(p) = 1 - \frac{\theta^3 e^{-\theta x}}{2\beta(1 + \theta) - \theta} \left[(\beta + \beta\theta - \theta) \frac{(x^2 \theta^2 + 2\theta x + 2)}{\theta^3} + \frac{x\theta + 1}{\theta^2} \right]$$

Où $x = F^{-1}(p)$ avec $F(\cdot)$ donnée par (2.20).

2.11.4 Statistiques d'ordres extrêmes

Si X_1, \dots, X_n un échantillon de n variables aléatoires qui suivent la distribution de Gamma **Lindley** et si $\bar{X} = (X_1 + \dots + X_n)/n$ représente la moyenne d'échantillon alors

par le théorème central limite $\frac{\sqrt{n}(\bar{X}-\mathbf{E}(X))}{\sqrt{\mathbf{Var}(X)}}$ se rapproche de la distribution normale standard quand $n \rightarrow \infty$.

Théorème 2.8

On va étudier la loi asymptotique des valeurs extrêmes

$M_n = \max(X_1, \dots, X_n)$ et $m_n = \min(X_1, \dots, X_n)$, Pour la fonction de répartition définie dans (2.20), on constate que

$$\lim_{t \rightarrow \infty} \frac{1 - F(t+x)}{1 - F(t)} = \exp(-\theta x).$$

et

$$\lim_{t \rightarrow 0} \frac{F(tx)}{F(t)} = x.$$

Ainsi, il résulte du **théorème 1.3** (théorème 1.6.2 dans **Leadbetter et al.** [21]) qu'il doit y avoir normalisation constantes $a_n > 0, b_n, c_n > 0$ et d_n de telle sorte que :

$$Pr\{a_n(M_n - b_n) \leq x\} \rightarrow \exp(-\exp(-\theta x)).$$

et

$$Pr\{c_n(m_n - d_n) \leq x\} \rightarrow 1 - \exp(-x). \quad (2.25)$$

Comme $n \rightarrow \infty$. La forme des constantes de normalisation peut également être déterminée. Par exemple, en utilisant le **Corollaire 1.1** (Corollaire 1.6.3 de **Leadbetter et al.** [21].), on peut voir que $a_n = 1$ et $b_n = F^{-1}(1 - 1/n)$ avec $F(\cdot)$ donnée par (2.20).

2.11.5 Fonction Quantile de la distribution de Gamma Lindley

Dans le théorème suivant, on donne une expression explicite de Q_X en fonction de la fonction W de **Lambert**.

Théorème 2.9

Pour tout $\theta > 0, \beta > \frac{\theta}{1+\theta}$, la fonction quantile de X qui suit la distribution de Gamma Lindley est :

$$Q_X(u) = -\frac{\beta(1+\theta)}{\theta(\beta(1+\theta)-\theta)} - \frac{1}{\theta} W_{-1} \left(\frac{\beta(1+\theta)(y-1)}{\beta(1+\theta)-\theta} e^{-\frac{\beta(1+\theta)}{\beta(1+\theta)-\theta}} \right), \quad 0 < u < 1. \quad (2.26)$$

Où W_{-1} désigne la branche négative de la fonction W de Lambert.

Preuve. Voir Théorème 1 (Zeghdoudi et Nedjar (2016))[11] ■.

2.11.6 Estimation

Estimation par la méthode des moments (MoM)

Étant donné un échantillon aléatoire X_1, \dots, X_n , de la distribution de Gamma Lindley (2.20), en utilisant le premier moment m et la variance s^2 on a :

$$m = \frac{2\beta(1+\theta)-\theta}{\theta\beta(1+\theta)}, s^2 = \frac{-(-2\beta\theta+\theta)^2 + (2+6\theta)\beta^2 - 2\beta(\beta\theta-3\beta\theta^2+2\theta^2)}{\beta^2\theta^2(1+\theta)^2}.$$

On résout ce système non linéaire pour tous $s > 0, m > 0$, pour trouver les estimateurs des moments $\hat{\theta}_{MoM}$ et $\hat{\beta}_{MoM}$ de θ et β respectivement comme suit :

$$\hat{\theta}_{MoM} = \frac{1}{s^2 + m^2} \left(2m + \sqrt{2}\sqrt{-s^2 + m^2} \right) \quad \text{et} \quad \hat{\beta}_{MoM} = \frac{\hat{\theta}}{(1 + \hat{\theta}) \left(2 - \hat{\theta}m \right)}. \quad (2.27)$$

L'estimation par la méthode du maximum de vraisemblance

Soient $X_i \sim GaL(\theta, \beta), i = \overline{1, n}$, n variables aléatoires. La fonction de logvraisemblance est :

$$\ln l(x_i; \theta, \beta) = 2n \ln \theta - n \ln \beta - n \ln(\theta + 1) + \sum_{i=1}^n \ln((\beta + \theta\beta - \theta)x_i + 1) - \theta \sum_{i=1}^n x_i. \quad (2.28)$$

Les dérivées de $\ln l(x_i; \theta, \beta)$ par rapport à θ et β sont : Les deux équations (2.42) et (2.43) ne peut pas résoudre directement, on doit utiliser la méthode de **Fisher Scoring**,

on a :

$$\begin{bmatrix} \frac{\partial^2 Lnl(x_i; \beta, \theta)}{\partial \theta^2} & \frac{\partial^2 Lnl(x_i; \beta, \theta)}{\partial \theta \partial \beta} \\ \frac{\partial^2 Lnl(x_i; \beta, \theta)}{\partial \beta \partial \theta} & \frac{\partial^2 Lnl(x_i; \beta, \theta)}{\partial \beta^2} \end{bmatrix}_{\substack{\hat{\theta} = \theta_0 \\ \hat{\beta} = \beta_0}} \begin{bmatrix} \hat{\theta} - \theta_0 \\ \hat{\beta} - \beta_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial Lnl(x_i; \beta, \theta)}{\partial \theta} \\ \frac{\partial Lnl(x_i; \beta, \theta)}{\partial \beta} \end{bmatrix}_{\substack{\hat{\theta} = \theta_0 \\ \hat{\beta} = \beta_0}} \quad (2.29)$$

L'équation (2.29) peut se résoudre de façon itérative où θ_0, β_0 sont les valeurs initiales de θ, β .

Les estimateurs dans le cas de l'observation unique (disons x_i) pour le vecteur de paramètres (θ, β) peut être écrit comme suit :

$$\begin{cases} \hat{\theta}_{MV} = \frac{1}{x} \\ \hat{\beta}_{MV} = \frac{1}{1+x} \end{cases}. \quad (2.30)$$

et

$$\begin{cases} E(\hat{\theta}_{MV}) = \frac{2\beta(1+\theta)-\theta}{\theta\beta(1+\theta)} = m \\ E(\hat{\beta}_{MV}) = \frac{e^\theta}{\beta(\theta+1)} (2\beta + \beta\theta - \theta - \beta\theta^2 + \theta^2). \end{cases} \quad (2.31)$$

2.12 Distribution de Size Biased Lindley Distribution et ses propriétés :

Les distributions biaisées par la taille apparaissent naturellement dans plusieurs contextes forestiers et écologiques. Les relations de pouvoir simples (par exemple, la surface terrière et le diamètre à hauteur de poitrine) entre les variables sont l'un de ces domaines d'intérêt découlant de la perspective de la modélisation. Un autre échantillon, un échantillon probabiliste proportionnel à la taille (PPT), se trouve dans les méthodes les plus largement utilisées pour échantillonner les matériaux debout ou morts et tombés dans la forêt. Il est souvent souhaitable ou nécessaire d'estimer une valeur paramétrique. modèle de densité de probabilité basé sur des données biaisées par la taille. Les méthodes traditionnelles d'égalité de probabilité peuvent ne pas être appropriées, ou peuvent être moins efficaces dans de telles circonstances, et il est préférable de procéder à une estimation en utilisant une théorie biaisée par la taille.

Les distributions biaisées par la taille apparaissent naturellement dans plusieurs contextes forestiers et écologiques. Simple relations de pouvoir (par exemple, la surface terrière et le diamètre à hauteur de poitrine) entre les variables sont une ce domaine d'intérêt découlant d'une perspective de modélisation.

Les distributions biaisées par la taille constituent un cas particulier de la forme plus générale connue sous le nom de distributions pondérées. Introduit pour la première fois par Fisher (1934) pour modéliser la détermination biais, ensuite les distributions pondérées ont été formalisées dans une théorie unificatrice par Rao (1965). De telles distributions apparaissent naturellement dans la pratique lorsque les observations d'un échantillon sont enregistrées avec une probabilité inégale, telle que de probabilité proportionnelle à dessins de taille (PPS). En bref, si la variable aléatoire X a une densité $f(x; \theta)$, avec paramètres inconnus θ , la densité pondérée correspondante est de la forme Size-biased distributions de forme :

$$f^\alpha(x, \theta) = \frac{w(x)f(x; \theta)}{E(w(X))}$$

où $w(x)$ est une fonction de pondération non négative telle que $E(w(X))$ existe. Un cas particulier des distributions pondérées, les distributions biaisées par la taille est proposé par Rao (1965) lorsque la fonction pondérée est de la forme $w(x) = x^\beta$, appelée distribution d'ordre biaisé par la taille β , où $\beta = 1$ ou $\beta = 2$, appelé longueur. respectivement à base de zones et à zones biaisées. Par conséquent, la densité de la distribution biaisée en longueur est défini par

$$f_L = \frac{xf(x, \theta)}{E(X)}, \quad -\infty < x < \infty,$$

la densité de La distribution de Lindley sera brièvement discutée

Dans cette section, on introduit la distribution de taille biaisée Lindley et étudier

ses propriétés de base. Plusieurs propriétés de cette distribution telles que les moments, la mesure de l'asymétrie, le kurtosis, la fonction génératrice de moments, la fonction génératrice de caractéristiques, le coefficient de la variation, de la fonction de survie et de la fonction de risque qui sont dérivés pour comprendre plus brièvement la structure de la distribution proposée. De plus, on converti la fonction de densité de la distribution Lindley (2.1) en son équivalent en fonction de taille (0.1). En suite, on trouve la fonction de densité de Lindley Biaisée sous la forme suivante :

$$f_{SBL}(x, \theta) = \frac{\theta^3(1+x)e^{-\theta x}}{2+\theta}, \quad x > 0. \quad (2.32)$$

Et la fonction de distribution cumulative du SBL est :

$$F_{SBL}(x, \theta) = \frac{2 - 2e^{-\theta x} - (2 + \theta x) + \theta(1 - e^{-\theta x}(1 + x\theta))}{2 + \theta}, \quad x > 0 \quad (2.33)$$

2.12.1 Fonction de survie et fonction de hasard

On a la fonction de survie

$$S_{SBL}(x) = \frac{e^{-\theta x}}{2 + \theta} [\theta + 2 - x\theta^2 + 2x\theta + x^2\theta^2], \quad (2.34)$$

et la fonction de hasard

$$h(x) = \frac{f_{SBL}(x)}{1 - F_{SBL}(x)} = \frac{\theta^3(1+x)}{\theta + 2 - x\theta^2 + 2x\theta + x^2\theta^2}. \quad (2.35)$$

2.12.2 Moments et mesures connexes

Le moment d'ordre k de la distribution de Size Baised **Lindley** est :

$$m_k = E(X^k) = \frac{\theta^2}{\theta^{k+2}} [(k+1)!(\theta+k+2)] \quad (2.36)$$

d'où, on a :

$$\mu_1' = \frac{(\theta+3)}{\theta(\theta+2)}, \mu_2' = \frac{6(\theta+4)}{\theta^2(\theta+2)}, \mu_3' = \frac{24(\theta+5)}{\theta^3(\theta+2)}, \mu_4' = \frac{120(\theta+6)}{\theta^4(\theta+2)}. \quad (2.37)$$

Le moment centré d'ordre k de la distribution de Size Baised **Lindley** est défini par :

$$\mu_k = \mathbf{E} \{(X - \mu)^r\} = \sum_{r=0}^k \binom{k}{r} \mu_k' (-\mu)^{k-r}.$$

D'où, on a

$$\begin{aligned} \mu_1 &= \frac{2(\theta+3)}{\theta(\theta+2)}, \mu_2 = \frac{2(\theta^2+6\theta+6)}{\theta^2(\theta+2)^2} = \sigma^2, \\ \mu_3 &= \frac{4(\theta^3+9\theta^2+18\theta+12)}{\theta^3(\theta+2)^3}, \mu_4 = \frac{24(\theta^4+12\theta^3+42\theta^2+60\theta+30)}{\theta^4(\theta+2)^4}. \end{aligned}$$

SBLD	μ_1	μ_2	μ_3	μ_4	Std. Dev
$\theta = 0.1$	29.52381	299.7732	5999.784	449591.7	17.31396
$\theta = 0.5$	5.6	11.84	47.872	708.4032	3.44093
$\theta = 0.9$	2.988506	3.584798	8.148448	65.90233	1.297429
$\theta = 1.3$	2.004662	1.683321	2.675344	14.75241	1.297429
$\theta = 1.7$	1.494436	0.9650163	1.181765	4.916901	0.9823524
$\theta = 2.1$	1.184669	0.6207837	0.6188595	1.025826	0.7878983
$\theta = 2.5$	0.9777778	0.4306173	0.3620521	1.002462	0.6562144
$\theta = 2.9$	0.8304011	0.3150689	0.2290128	0.5418929	0.56131
$\theta = 3.3$	0.7204117	0.2398822	0.1535249	0.3168071	0.4897777

Tableau 2.3. Moments et écart type pour différentes valeurs du paramètre θ .

Le coefficient de variation (γ), le coefficient de dissymétrie *Skewness* ($\sqrt{\beta_1}$) et le coefficient d'aplatissement *Kurtosis* (β_2) sont :

$$\gamma = \frac{\sqrt{2(\theta^2 + 6\theta + 6)}}{2(\theta + 3)},$$

$$Skewness(\sqrt{\beta_1}) = \frac{4(\theta^3 + 9\theta^2 + 180\theta + 12)}{(2(\theta^2 + 6\theta + 6))^{\frac{3}{2}}},$$

$$Kurtosis(\beta_2) = \frac{24(\theta^4 + 12\theta^3 + 42\theta^2 + 60\theta + 30)}{(2(\theta^2 + 6\theta + 6))^2}.$$

Table 2.4. *Skewness, kurtosis et le coefficient de variation de SBLD pour certaines valeurs du paramètre θ*

θ	Skewness pour SBLD	Kurtosis pour SBLD	CV pour SBLD
0.1	1.155969	5.003023	0.5864406
0.5	1.175044	5.053324	0.6144518
0.9	1.200544	5.128277	0.6335461
1.3	1.224981	5.206302	0.6472056
1.7	1.246606	5.279858	0.6573401
2.1	1.265265	5.34658	0.6650788
2.5	1.28125	5.406111	0.6711283
2.9	1.294946	5.458867	0.6759504
3.3	1.306718	5.505525	0.6798582

Chapitre 3

Distribution de Size Baised

Gamma Lindley (SBGaL)

immédiate propriétés

3.1 Introduction

Dans ce chapitre, on propose une nouvelle extension de la distribution Gamma Lindley appelée la distribution de Size Baised Gamma Lindley (SBGaL). Plusieurs propriétés de cette distribution telles que la méthode du moment, l'estimation du maximum de vraisemblance et la distribution limite des statistiques d'ordre extrême sont établies. Une étude de simulation est réalisée pour examiner le biais et l'erreur quadratique moyenne des estimateurs du maximum de vraisemblance des paramètres.

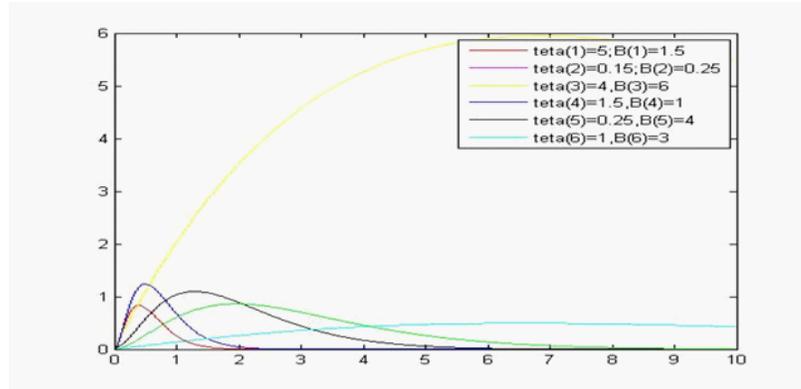


FIG. 3.1: **Fig 3.1.** *Présentation graphique de la fonction de densité pour quelques valeurs de θ .*

3.2 Distribution Size Baised Gamma-Lindley (SB-GaLD) et certaines propriétés

En faisant une transformation de la fonction de densité de la distribution Gamma Lindley (2.18) en utilisant la formule (0.1), on trouve la fonction de densité de Size Baised Gamma-Lindley :

$$f_{SBGaL}(x; \theta; \beta) = \begin{cases} \frac{\theta^3((\beta + \beta\theta - \theta)x^2 + x)e^{-\theta x}}{2\beta(1 + \theta) - \theta} & \text{si } x, \theta > 0 \text{ et } \beta > \frac{\theta}{1 + \theta} \\ 0 & \text{si non} \end{cases} \quad (3.1)$$

Et la fonction de distribution cumulative du SBGaL est :

$$F_{SBGaL}(x) = 1 - \left(\frac{\theta^2(\beta + \theta\beta - \theta)}{2\beta(1 + \theta) - \theta} x^2 + \theta x + 1 \right) e^{-\theta x} \text{ si } x, \theta > 0 \text{ et } \beta > \frac{\theta}{1 + \theta} \quad (3.2)$$

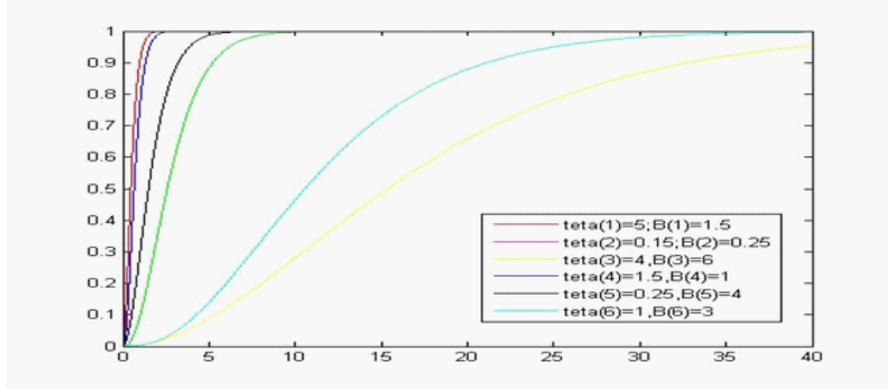


FIG. 3.2: **.2.** *Présentation graphique de la fonction de répartition pour quelques valeurs de θ .*

Par conséquent, le mode de SBGaL est donné par :

$$Mode(X) = \begin{cases} \frac{(\beta + \theta\beta - \frac{3}{2}\theta) + \sqrt{(\beta + \theta\beta - \theta)^2 + \frac{\theta^2}{4}}}{\theta(\beta + \theta\beta - \theta)}, & \text{si } \beta \in \left[\frac{\theta}{1+\theta}, \infty\right[\\ 0 & \text{si non} \end{cases} \quad (3.3)$$

3.2.1 Fonction de survie et la La fonction de hasard

Les fonctions de survie et de hasard sont :

$$S_{SBGaL}(x) = 1 - F_{SBGaL}(x) = \left(\frac{\theta^2(\beta + \theta\beta - \theta)}{2\beta(1 + \theta) - \theta} x^2 + \theta \left(x + \frac{1}{\theta} \right) \right) e^{-\theta x}, \quad (3.4)$$

et

$$h(x) = \frac{f_{SBGaL}(x)}{1 - F_{SBGaL}(x)} = \frac{\theta^2((\beta + \theta\beta - \theta)x^2 + x)}{\theta(\beta + \theta\beta - \theta)x^2 + (2\beta(1 + \theta))(x + \frac{1}{\theta})}. \quad (3.5)$$

3.3 Statistiques Moments, courbe de Lorenz et ordre extrême

3.3.1 Moments et mesures connexes

Le kième moment sur l'origine du SBGaL est :

$$m_k = E(X^k) = 3 \frac{6^{k-1}}{\theta^k} - \frac{6^{k-1}}{\theta^{k-1}(2\beta(1+\theta) - \theta)} - \frac{6}{25} \frac{(-5k + 6^k - 1)}{\theta^k} \quad (3.6)$$

Proposition 3.1

Soit X_1, X_2, \dots, X_n n variables aléatoires indépendantes de la distribution SBGaL(θ, β).

Puis la fonction génératrice du moment (mgf) de $S = \sum_{i=1}^n X_i$, est donnée par :

$$M_X(t) = E(e^{tX}) = \theta^3 \left(\frac{-1}{(t-\theta)^3} - \frac{t}{(2\beta(1+\theta\beta) - \theta)(t-\theta)^3} \right) \quad (3.7)$$

et

$$M_S(t) = \theta^{3n} \left(\frac{-1}{(t-\theta)^3} - \frac{t}{(2\beta(1+\theta\beta) - \theta)(t-\theta)^3} \right)^n \quad (3.8)$$

Preuve. Soit X_1, X_2, \dots, X_n variables aléatoires indépendantes, nous avons

$$M_S(t) = E(e^{tS}) = M_X(t)^n = \theta^{3n} \left(\frac{-1}{(t-\theta)^3} - \frac{t}{(2\beta(1+\theta\beta) - \theta)(t-\theta)^3} \right)^n$$

Remarque 3.1

La fonction génératrice de moment pour X et S existes ($E(e^{tX}) < \infty$) que si $t < \theta$.

Corollaire 3.1

Soit $X \sim SBGaL(\theta, \beta)$ alors la moyenne et la variance de X sont :

$$E(X) = \frac{2}{\theta} - \frac{1}{2\beta(1+\theta) - \theta}, E(X^2) = \frac{6\beta(1+\theta) - 4\theta}{\theta(2\beta(1+\theta) - \theta)} \quad (3.9)$$

$$Var(X) = \frac{3}{\theta^2} - \frac{1}{(2\beta(1+\theta) - \theta)^2} \quad (3.10)$$

Preuve.

On a

$$E(X) = M'(t=0) = \frac{2}{\theta} - \frac{1}{2\beta(1+\theta) - \theta}, E(X^2) = M''(t=0) = \frac{6\beta(1+\theta) - 4\theta}{\theta(2\beta(1+\theta) - \theta)} \quad (3.11)$$

et la variance

$$Var(X) = E(X^2) - E(X)^2 = \frac{3}{\theta^2} - \frac{1}{(2\beta(1+\theta) - \theta)^2} \quad (3.12)$$

Remarque 3.2

La relation entre les moments est la suivante :

$$\frac{\theta}{6}m_k = m_{k-1} - \frac{k-1}{\theta^k}, k = 0, 1, 2, \dots \quad (3.13)$$

Cette égalité a été vérifiée avec démonstration par récurrence.

Théorème 3.1

Soit $X \sim SBGaL(\theta, \beta)$, $M = mode(X)$, $me = median(X)$, $\mu = E(X)$, et $\theta > 0, \beta > \frac{\theta}{1+\theta}$. Alors

$$M < me < \mu \quad (3.14)$$

Preuve.

Soit

$$F_{SBGaL}(M) = 1 - \left(\frac{[(\beta + \theta\beta - \frac{3}{2}\theta) + (\sqrt{(\beta + \theta\beta - \theta)^2 - \frac{\theta^2}{4}})]^2}{(2\beta(1+\theta) - \theta)(\beta + \theta\beta - \theta)} + \frac{(\beta + \theta\beta - \frac{3}{2}\theta) + \sqrt{(\beta + \theta\beta - \theta)^2 - \frac{\theta^2}{4}}}{(\beta + \theta\beta - \theta)} + 1 \right) e^{-\frac{(\beta + \theta\beta - \frac{3}{2}\theta) + \sqrt{(\beta + \theta\beta - \theta)^2 - \frac{\theta^2}{4}}}{(\beta + \theta\beta - \theta)}},$$

$$F_{SBGaL}(\mu) = 1 - \left(\frac{(\beta + \theta\beta - \theta)(6\beta(1+\theta) - 4\theta)^2}{(2\beta(1+\theta) - \theta)^3} + \frac{6\beta(1+\theta) - 4\theta}{2\beta(1+\theta) - \theta} + 1 \right) e^{-\frac{6\beta(1+\theta) - 4\theta}{2\beta(1+\theta) - \theta}}.$$

Notez que

$$\frac{6\beta(1+\theta) - 4\theta}{2\beta(1+\theta) - \theta} > 1, \quad \frac{[(\beta + \theta\beta - \frac{3}{2}\theta) + \sqrt{(\beta + \theta\beta - \theta)^2 - \frac{\theta^2}{4}}]}{(\beta + \theta\beta - \theta)} < 2 \text{ pour } \beta > \frac{\theta}{1+\theta} \quad (3.15)$$

Il est facile de vérifier que

$$F(M) < F(me) < F(\mu),$$

et finalement, on obtient

$$M < me < \mu.$$

3.3.2 Courbe de Lorenz

La courbe de Lorenz pour une variable aléatoire positive est définie comme le graphique du ratio

$$L(F(x)) = \frac{E(X \mid X \leq x)}{E(X)}.$$

Contre $F(x)$ avec les propriétés $L(p) = p, L(1) = 1$. Si X représente un revenu annuel, $L(p)$ est la proportion du revenu total revenant aux individus ayant les 100 p % revenus les plus bas. Si tous les individus gagnent le même revenu, alors $L(p) = p$ pour tout p . La zone située entre la ligne et la courbe de Lorenz peut être considérée comme une mesure de l'inégalité des revenus, ou plus généralement de la variabilité de la variabilité de x . see [27] et de [2] pour une analyse approfondie de Lorenz. courbes. Pour la distribution exponentielle, il est bien connu que la courbe de Lorenz est donnée par :

$$L(p) = p\{p + (1 - p) \log(1 - p)\}$$

Pour la distribution en (3.2),

$$E(X \mid X \leq x)F(x) = \frac{6\beta(1+\theta)-4\theta}{\theta(2\beta(1+\theta)-\theta)} - \frac{e^{-\theta x}}{2\beta(1+\theta)-\theta} [(\beta+\theta\beta-\theta)\left(\frac{6}{\theta}+6x+3\theta x^2+\theta^2 x^3\right)+2+2\theta x+\theta^2 x^2]$$

Ainsi, en (3.2) nous obtenons la courbe de Lorenz pour la distribution Size Biased Gamma Lindley comme

$$L(p) = 1 - \frac{(2\beta(1+\theta)-\theta)\left(\frac{\theta^2(\beta+\theta\beta-\theta)}{2\beta(1+\theta)-\theta}x^2+\theta x+1\right)\left[(\beta+\theta\beta-\theta)\left(\frac{6}{\theta}+6x+3\theta x^2+\theta^2 x^3\right)+2+2\theta x+\theta^2 x^2\right]e^{-\theta x}}{(6\beta(1+\theta)-4\theta)\left[\frac{\theta(\beta+\theta\beta-\theta)}{2\beta(1+\theta)-\theta}x^2+x+\frac{1}{\theta}\right]} \quad (3.16)$$

où $x = F^{-1}(p)$ avec $F(\cdot)$ donné par (3.2).

3.3.3 Statistiques de commande extrêmes

Si X_1, \dots, X_n est un échantillon aléatoire de (3.1) et si $\bar{X} = \frac{X_1 + \dots + X_n}{n}$ désigne l'échantillon signifie alors par le théorème de la limite centrale habituelle $\frac{\sqrt{n}(\bar{X} - \mathbf{E}(X))}{\sqrt{\text{Var}(X)}}$ se rapproche de la distribution normale standard comme $n \rightarrow \infty$. Parfois, on serait intéressé par l'asymptotique des valeurs extrêmes $M_n = \max(X_1, \dots, X_n)$ et $m_n = \min(X_1, \dots, X_n)$. Pour le c.d.f. en (3.5), on peut voir que

$$\lim_{t \rightarrow \infty} \frac{1 - F(t+x)}{1 - F(t)} = e^{-\theta x} \quad (3.17)$$

$$\lim_{t \rightarrow 0} \frac{F(tx)}{F(t)} = x \quad (3.18)$$

Ainsi, il ressort du théorème 1.6.2 de [21] doit être des constantes de normalisation $a_n, b_n, c_n > 0$ et d_n telque :

$$P(a_n(M_n - b_n) \leq x) \xrightarrow[n \rightarrow \infty]{} e^{e^{-\theta x}} \quad (3.19)$$

La forme des constantes normatives peut également être déterminée. Par exemple, en utilisant le corollaire 1.6.3 dans [21], on peut voir que $a_n = 1$ et $b_n = F^{-1}\left(1 - \frac{1}{n}\right)$ avec $F(\cdot)$ donné par (3.5).

3.4 Estimation des paramètres

3.4.1 Estimation du maximum de vraisemblance (ML)

Dans cette section, nous discuterons de l'estimation ponctuelle sur les paramètres qui indexent la fonction. Laissez la fonction log-vraisemblance d'une seule observation (disons) pour le vecteur de paramètre s'écrivant comme suit :

$$\ln l(x_i; \theta, \beta) = 3 \ln \theta + \ln((\beta + \theta\beta - \theta)x_i + 1) + \ln x_i - \theta x_i - \ln(2\beta(1 + \theta) - \theta) \quad (3.20)$$

Les dérivées de $\ln l(x_i; \theta)$ par rapport à θ et β sont :

$$\frac{d \ln l(x_i; \theta, \beta)}{d\theta} = \frac{3}{\theta} - \frac{2\beta - 1}{2\beta(1 + \theta) - \theta} - x_i - \frac{(\beta - 1)x_i}{(\beta + \theta\beta - \theta)x_i + 1} \quad (3.21)$$

$$\frac{d \ln l(x_i; \theta, \beta)}{d\beta} = -\frac{2(1 + \theta)}{2\beta(1 + \theta) - \theta} + \frac{(1 + \theta)x_i}{(\beta + \theta\beta - \theta)x_i + 1} \quad (3.22)$$

L'estimateur du maximum de vraisemblance $\hat{\theta}$ de θ et $\hat{\beta}$ de β est obtenu en résolvant les équations non linéaires (3.21) et (3.22) fournies

$$\begin{aligned} \hat{\theta} &= \frac{2}{x} \\ \hat{\beta} &= \frac{2}{x + 2} \end{aligned} \quad (3.23)$$

et

$$\begin{aligned}
E(\widehat{\theta}) &= \frac{2\theta\beta(1+\theta)}{2\beta(1+\theta)-\theta} \\
E(\widehat{\beta}) &= \frac{2\theta((\beta+\theta\beta-\theta)(1-2\theta)+\theta)+2\theta^3(-4(\beta+\theta\beta-\theta)+2)e^{2\theta}E(-2\theta)}{2\beta(1+\theta)-\theta}
\end{aligned} \tag{3.24}$$

3.4.2 Moments estimés

En utilisant le premier moment m et m_2 second moment sur la distribution SBGaL, nous avons

$$\begin{aligned}
m &= \frac{3}{\theta} - \frac{1}{2\beta(1+\theta)-\theta} \\
m_2 &= \frac{12}{\theta^2} - \frac{6}{\theta(2\beta(1+\theta)-\theta)}
\end{aligned} \tag{3.25}$$

Où $m_2 = s^2 + m^2$ et s^2 est la variance. Nous résolvons ce système non linéaire et nous trouvons le couple $(\widehat{\theta}, \widehat{\beta})$, Où $(\widehat{\theta}, \widehat{\beta}) > 0$ également $s > 0, m > 0$, la résolution du système non linéaire (3.25) donne :

$$m_2\theta^2 - 6m\theta + 6 = 0 \text{ et } \beta = \frac{4\theta - \theta^2 m}{2(1+\theta)(3-\theta m)} \tag{3.26}$$

Il est facile de vérifier que la solution de

$$\widehat{\theta} = \frac{3m + \sqrt{3}\sqrt{m^2 - 2s^2}}{m_2} \tag{3.27}$$

et

$$\hat{\beta} = \frac{4\hat{\theta} - \hat{\theta}^2 m}{2(1 + \hat{\theta})(3 - \hat{\theta}m)} \quad (3.28)$$

3.5 Simulation

Dans cette section, nous étudions le comportement des estimateurs MV pour un échantillon de taille (n) finie. Une étude de simulation comporte les étapes suivantes est effectuée pour chaque triple (n, θ, β) , où $\theta = 0.1, 0.5, 1, 3$, $\beta = 0.1, 0.5, 0.75, 1, 6$ et $n = 10, 30, 50$.

choisir les valeurs initiales de θ_0, β_0 pour les éléments correspondants du vecteur de paramètres $\Theta = (\theta; \beta)$ pour spécifier la distribution SBGaL.

Choisir les valeurs initiales θ_0, β_0 pour les éléments correspondants au vecteur de paramètres $\Theta = (\theta, \beta)$ pour spécifier la distribution SBGaL ;

- Choisir la taille de l'échantillon n ;
- Générer N échantillons indépendants de taille n à partir de $SBGaL(\theta, \beta)$;
- Calculer les estimations $\hat{\Theta}_{MV}$ de Θ pour chacun des N échantillons ;
- Calculer la moyenne des estimateurs obtenus pour N échantillons,

$$\text{biais moyen}(\Theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\Theta}_i - \Theta_0)$$

et l'erreur quadratique moyenne

$$MSE(\Theta) = \frac{1}{N} \sum_{i=1}^N \left(\hat{\Theta}_i - \Theta_0 \right)^2.$$

Tableau 3.1 : *Biais moyen des estimations simulées* $\theta = 1, \beta = 6$, $\theta = 1, \beta = 0.1$, $\theta = 1, \beta = 0.75$

	$\theta = 1, \beta = 6$		$\theta = 1, \beta = 0.6$		$\theta = 1, \beta = 0.75$	
	<i>biais</i> (θ)	<i>biais</i> (β)	<i>biais</i> (θ)	<i>biais</i> (β)	<i>biais</i> (θ)	<i>biais</i> (β)
$n = 10$	0.0043	-0.5549	0.1666	0.0162	0.0500	-0.0250
$n = 30$	0.0014	-0.1849	5.5556×10^{-2}	0.0054	1.6666×10^{-2}	-0.0020
$n = 50$	0.0008	-0.1109	3.3334×10^{-2}	0.0032	0.0100	-0.0012
	$\theta = 0.1, \beta = 1$		$\theta = 0.5, \beta = 1$		$\theta = 3, \beta = 1$	
	<i>biais</i> (θ)	<i>biais</i> (β)	<i>biais</i> (θ)	<i>biais</i> (β)	<i>biais</i> (θ)	<i>biais</i> (β)
$n = 10$	0.0004	-0.0911	0.0100	-0.0680	0.1800	-0.0858
$n = 30$	1.5800×10^{-4}	-0.0303	3.3333×10^{-2}	-0.0226	0.0600	-0.0286
$n = 50$	0.0096×10^{-2}	-0.0182	0.0020	-0.0136	0.0360	-0.0171
	$\theta = 3, \beta = 0.8$		$\theta = 0.5, \beta = 0.5$		$\theta = 0.1, \beta = 0.5$	
	<i>biais</i> (θ)	<i>biais</i> (β)	<i>biais</i> (θ)	<i>biais</i> (β)	<i>biais</i> (θ)	<i>biais</i> (β)
$n = 10$	0.9000	0.8286	0.0250	-0.0149	0.0010	-0.0408
$n = 30$	0.3000	0.2762	8.3333×10^{-3}	-0.0049	3.3333×10^{-4}	-0.0136
$n = 50$	0.1800	0.1657	0.0050	-0.0029	0.0002	-0.0081

Le tableau 3.1 montre que θ est biaisé positivement et que le $biais(\theta) \rightarrow 0, \theta \rightarrow 0$,

et β est biaisé négativement si $\beta > 0,5$ et le $biais(\beta) \rightarrow 0, \beta \rightarrow 0$.

Tableau 3.2 : l'erreur quadratique moyenne MSE

	$\theta = 1, \beta = 6$		$\theta = 1, \beta = 0.5$		$\theta = 1, \beta = 0.75$	
	$MSE(\theta)$	$MSE(\beta)$	$MSE(\theta)$	$MSE(\beta)$	$MSE(\theta)$	$MSE(\beta)$
$n = 10$	1.8922×10^{-4}	3.080	0.2777	0.0026	0.0250	0.0062
$n = 30$	6.3075×10^{-5}	1.0267	0.0925	8.8551×10^{-4}	8.3333×10^{-3}	0.0020
$n = 50$	3.7845×10^{-5}	0.6160	0.0555	0.0005	0.0050	0.0012
	$\theta = 0.1, \beta = 1$		$\theta = 0.5, \beta = 1$		$\theta = 3, \beta = 1$	
	$MSE(\theta)$	$MSE(\beta)$	$MSE(\theta)$	$MSE(\beta)$	$MSE(\theta)$	$MSE(\beta)$
$n = 10$	2.3×10^{-6}	0.0830	0.0010	0.0463	0.3240	0.0736
$n = 30$	7.6800×10^{-7}	0.0276	3.3×10^{-4}	0.0154	0.1080	0.0245
$n = 50$	4.6080×10^{-7}	0.0166	0.0002	0.0092	0.0648	0.0147
	$\theta = 3, \beta = 0.8$		$\theta = 0.5, \beta = 0.5$		$\theta = 0.1, \beta = 0.5$	
	$MSE(\theta)$	$MSE(\beta)$	$MSE(\theta)$	$MSE(\beta)$	$MSE(\theta)$	$MSE(\beta)$
$n = 10$	8.1	6.8669	6.25×10^{-3}	0.0022	$0,1 \times 10^{-4}$	0.0166
$n = 30$	2.7	2.2889	2.08×10^{-3}	0.0007	3.3×10^{-6}	0.0055
$n = 50$	1.62	1.3733	0.0012	0.0004	0.2×10^{-5}	0.0033

Le tableau 3.2 montre que $MSE(\theta)$ et $MSE(\beta) \rightarrow 0$, pour $\theta \rightarrow 0$ et $n \rightarrow \infty$

3.5.1 Application aux ensembles de données réelles

Dans cette section, nous illustrons l'applicabilité de la distribution en considérant deux ensembles de données différents utilisés par des chercheurs différents. Nous avons également ajusté Lindley généralisé par H. Zakerzadah, A. Dolati (2010), quasi Lindley par R. Shanker, A. Mishra (2013), Two-Parameter Lindley de R. Shanker, S. Sharma (2013), Weibull et lognormal distributions.

Dans chacune de ces distributions, les paramètres sont estimés à l'aide de la méthode moment car elle est simple, facile à manipuler, exacte et, pour la comparaison, nous utilisons des valeurs de log-vraisemblance négatives ($-LL$), le critère d'information d'Akaike (AIC) et des informations bayésiennes. critère (BIC) qui sont définis par $-2LL + 2q$ et $-2LL + q \log(n)$, respectivement, où q est le nombre de paramètres estimés et n la taille de l'échantillon. Une autre statistique de test $K-S$ (Kolmogorov - Smirnov) définie comme $K-S = \sup |F_n(x) - F(x)|$ où $F_n(x)$ est une fonction de distribution empirique et $F(x)$ une fonction de distribution cumulative est calculée et affichée pour tous les ensembles de données.

Exemple 1

Nous considérons dans Lawless de J. F. Lawless (2003), pages 204 et 263, deux séries de données réelles. Le premier représente les temps de défaillance (mm) pour un échantillon de quinze composants électroniques dans une accélération test de durée de vie : 1.4, 5.1, 6.3, 10.8, 12.1, 18.5, 19.7, 22.2, 23, 30.6, 37.3, 46.3, 53.9, 59.8, 66.2.

Le second ensemble de données est le nombre de cycles jusqu'à la rupture pour 25 échantillons de fil de 100 cm, testés à un niveau de contrainte particulier : 15, 20, 38, 42, 61, 76, 86, 98, 121, 146, 149 , 157, 175, 176, 180, 180, 198, 220, 224, 251, 264, 282, 321, 325, 653.

Tableau 3.3 *Comparaison entre les distributions*

Données	Distribution	β	θ	γ	$-LL$	$K - S$	AIC	BIC
$n = 15$	Generalized Lindley	1.203	0.064	0.083	64.080	0.095	134.16	136.28
$n = 15$	GaL	1.129	0.684		64.015	0.094	132.03	133.45
$m = 27.546$	QLD	4 .016	-0.99		1504	0.93	3012	3013.4
$s = 20.059$	TwoPLD	0.0704	1.110			0.196		
	Gamma	1.442	0.052		64.197	0.102	132.39	133.81
	Weibull	1.306	0.034		64.026	0.450	132.05	133.47
	Lognormal	1.061	2.931		64.626	0.163	135.25	136.67
	SBLD		1.84		64.107	0.106	133.23	135.55
	SBGaL	1.295	0.84		63.52	0.089	132.02	133.33
Série2	Generalized Lindley	1.505	0.012	0.018	152.369	0.137	310.74	314.39
$n = 25$	GaL	1.086	0.010		152.132	0.129	308.26	310.7
$m = 178.32$	QLD	0.0107	8.514		1045.9	0.94	2131.8	2156.2
$s = 131.097$	TwoPLD	0.0107	0.125			0.232		
	Gamma	1.794	0.010		152.371	0.135	308.74	311.18
	Weibull	1.414	0.005		152.440	0.697	308.88	310.7
	Lognormal	0.891	4.880		154.092	0.155	312.18	314.62
	SBLD		0.080		155.12	0.159	313.56	315.78
	SBGaL	1.112	0.069		151.35	0.125	307.67	310.7

Selon le tableau 3, nous pouvons observer que la distribution gamma-Lindley bai-

sée par la taille fournit le plus petit $-LL$, AIC , $K-S$ et BIC valeurs comparées aux distributions de Lindley généralisées, Gamma Lindley, quasi-Lindley, distributions de Lindley à deux paramètres, Gamma, Weibul, Lognormales, biais de taille, et correspondent donc mieux aux données de tous les modèles considérés.

Conclusion et Perspectives

Ainsi, nous avons réussi à introduire une nouvelle distribution nommée "distribution de **Size Baised Gamma Lindly**" dont la distribution de Lindley basée sur la taille (SBL), est un cas particulier, a été introduite pour modéliser les données de comptage excluant structurellement les comptages nuls. L'estimation de ses paramètres a été discutée en utilisant la méthode du maximum de vraisemblance et la méthode des moments. Une étude de simulation est réalisée pour examiner le biais et l'erreur quadratique moyenne des estimateurs du maximum de vraisemblance des paramètres. Plusieurs applications du modèle à un ensemble de données réelles sont enfin présentées et comparées à l'ajustement obtenu par quelques autres paramètres bien connus. L'adéquation des ajustements a été évaluée en termes de valeurs AIC , de valeurs BIC et de courbes de densité. Nous pouvons montrer que la distribution gamma-Lindley basée sur la taille peut être utilisée assez efficacement pour analyser des données réelles et la science actuarielle.

Nous pourrions dans nos recherches futures proposer d'autres distributions à savoir :

- Inverse pseudo Lindley distribution.

- Size biased Zeghdoudi distribution
- Size biased pseudoLindley distribution .

Bibliographie

- [1] **Asgharzadeh Hassan A., Bakouch,S.,& Esmaeili ,L.** (2013). Pareto poisson-**Lindley** distribution and its application, *Journal of Applied Statistics*, pp. 1-18.
- [2] **C. Dagum**, Lorenz curve, in : S. Kotz, N.L. Johnson, C.B. Read (Eds.), in : *Encyclopedia of Statistical Sciences*, vol. 5,Wiley, New York, 1985,pp. 156–161.
- [3] **E. M. Lémeray**, Racines de quelques équations transcendantes. Intégration d'une équation aux différences mêlées. Racines imaginaires,*Nouvelles Ann. Math.* 16 (1897) 540–546.
- [4] **F. Lawless**, *Statistical Models and Methods for Lifetime Data*, Wiley, New York, 200
- [5] **Glaser, R. E.** (1980). Bathtub and related failure rate characterizations. *J. Amer. Statist. Assoc.*, 75, 667-672
- [6] **H. S. Bakouch, B. M. Al-Zahrani, A. A. Al-Shomrani, V.A. Marchi and F. Louzada (2012).** An extended **Lindley** distribution, *Journal of the Korean Statistical Society*, Vol **41**, 75-85.

- [7] **Hussain, E.** The Non-Linear Functions of Order Statistics and their Properties in Selected Probability Models (Doctoral dissertation, University of Karachi, Karachi) (2008).
- [8] **H. Zakerzadah, A. Dolati** (2010). Generalized **Lindley** distribution. *J. Math. Ext.* 3(2),pp. 13-25..
- [9] **H. Zeghdoudi, S. Nedjar** (2016a). Gamma **Lindley** distribution and its application. *Journal of Applied Probability and Statistics* Vol. **11**, N° **1**, 129-138
- [10] **H. Zeghdoudi, S. Nedjar** (2016b) A Pseudo **Lindley** distribution and its application., *J. Afrika Statistika* Vol 11 (1) page 923_932.
- [11] **H. Zeghdoudi, S. Nedjar** (2016c). On Gamma **Lindley** distribution :properties and simulation. *Journal of Computational and Applied Mathematics* 298, pp167-174.
- [12] **H. Zeghdoudi, S. Nedjar** (2017). A Poisson pseudo **Lindley** distribution and its application . *Journal of Probability and Statistical Sciences* . Vol. 15, N° 1.
- [13] **H. Zeghdoudi, S. Nedjar** (2017). On Pseudo **Lindley** Distribution : Properties and Applications . *Journal of New Trends in Mathematical Sciences*
- [14] **J. F. Lawless** , 2003. *Statistical models and methods for lifetime data.* Wiley, New York pp 204 - 263.
- [15] **Jodrá, P.** (2010). Computer generation of random variables with **Lindley** or

- Poisson **Lindley** distribution via the **Lambert** W function. Mathematics and Computers in Simulation (MATCOM), Vol. 81, issue 4, pages 851-859.
- [16] **Lindley**, D. V. (1958). Fiducial distributions and bayes theorem. Journal of the Royal Society, series B, 20, 102-107
- [17] M. E. Ghitany, B. Atieh, S. Nadarajah (2008a). **Lindley** distribution and its applications. *Math. Comput. Simulation* , **78**, pp. 493-506 .
- [18] M..E. Ghitany, D.K. Al-Mutairi, S. Nadarajah (2008b). Zero-truncated Poisson–**Lindley** distribution and its application, *Math. Comput.*
- [19] M..E.Ghitany.D.K Almutairi(2009) : Estimation methods for the discrete Poisson–**Lindley** distribution, Journal of Statistical Computation and Simulation, 79 :1, 1-9
- [20] M. H. Gail and J. L.Gastwirth (1978). A scale-free goodness of fit test for the exponential distribution based on the Lorenz curve. *Journal of the American Statistical Association* **73** 787–793.
- [21] M..R. Leadbetter, G. Lindgren, H. Rootz´en, Extremes and Related Properties of Random Sequences and Processes, Springer Verlag, New York,1987.
- [22] M. Sankaran (1970). The discrete Poisson-**Lindley** distribution. *Biometrics*, **26**, pp. 145-149.
- [23] M. Shaked, J.G. Shanthikumar, Stochastic Orders and Their Applications, Academic Press, New York, 1994..

- [24] R.Shanker and A. Mishra (2013). A quasi **Lindley** distribution. African Journal of Mathematics and Computer Science Research, Vol. 6(4), pp. 64-71.
- [25] R.Shanker, S.Sharma, R.Shanker. (2013). A two-parameter **Lindley** distribution for modeling waiting and survival times data. *Applied Mathematics*, vol. 4, pp. 363-368.
- [26] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, D.E. Knuth (1996). On the **Lambert** W function, *Adv. Comput. Math.* **5** 329–359.
- [27] Patil, G. P. and Rao, C. R. (1978) : Weighted distributions and size-biased sampling with applications to wild life populations and human families, *Biometrics*, 34, 179 - 189.
- [28] Gilbert Colletaz (2016) : Modèles de survie. Notes de cours p9.